In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

University of Alberta

The dominant MHC class I locus of the duck (*Anas platyrhynchos*) is adjacent to the polymorphic TAP2 locus

by

Christine Marie Mesa

A thesis
submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Physiology and Cell Biology

Department of Biological Sciences

Edmonton, Alberta

Fall 2003

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisisitons et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-87836-8
Our file Notre référence
ISBN: 0-612-87836-8

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Canada

University of Alberta

Library Release Form

Name of Author: Christine Marie Mesa

Title of Thesis: The dominant MHC class I locus of the duck (*Anas platyrhynchos*) is adjacent to the polymorphic TAP2 locus.

Degree: Master of Science

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

August 26, 2003

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled: "The dominant MHC class I locus of the duck (*Anas platyrhynchos*) is adjacent to the polymorphic TAP2 locus" submitted by Christine Marie Mesa in partial fulfillment of the requirements for the degree of Master of Science in Physiology and Cell Biology.

Dr. Katharine E. Magor

Dr. Cindy Paszkowski

Dr. John F. Elliott

August 15, 2003

Abstract

We examined the expression and linkage of the major histocompatibility complex (MHC) class I genes and the transporter associated with antigen processing 2 (TAP2) gene in the duck (*Anas platyrhynchos*). Southern blot analysis and amplification of genomic sequences suggest that ducks have at least four genes encoding MHC class I. We constructed and screened a primary cDNA library and identified four transcripts, with overwhelming expression of one. To investigate whether this was due to linkage of the dominant MHC class I locus with TAP2, we cloned the TAP2 cDNA. Amplification of the genomic region between TAP2 and class I by PCR showed the dominantly expressed MHC class I locus is tightly linked to TAP2. TAP2 is also polymorphic and amino acid differences between the TAP2 alleles may have functional consequences for what antigenic peptide can be transported and presented. The organization of the duck MHC may have functional consequences for their ability to eliminate viral pathogens.

Acknowledgements

Thank-you to all who have been instrumental in helping me complete this thesis. Special thanks go to my supervisor, Kathy Magor, who gave me this opportunity to learn and grow; you have my never-ending gratitude. Thanks to my committee members, other faculty, and support staff from the department of Biological Sciences who encouraged me and gave me guidance during my time here. I don't think I would have had the wherewithal if I hadn't known so many people were cheering me on. To my 'ultimate' friends who made sure I always had some fun, and to those that I shared the trials and tribulations of grad studies with; remember, if there is no time for beer, something is wrong. There should always be time for beer. Cheers! I cannot leave out my family, who supported me in everything I've done and let me follow my dreams; I owe everything to you. And most especially to my husband Adam, who has helped me believe in my dreams; without you, this would have never been done. Thank-you for believing in me.

Table of Contents

1 INTRODUCTION 1

1.1 Major Histocompatibility Complex (MHC) class I 1
1.1.1 Structure of MHC class I molecule 1
1.1.2 Assembly of MHC class I 3
1.1.3 Functions of MHC class I 7
1.1.4 Diversity in MHC class I 15

1.2 Transporter associated with antigen processing 2 (TAP2) 26
1.2.1 Structure and function 26

1.3 Organization of the MHC in vertebrates and functional consequences 29
1.3.1 Cartilaginous fish 31
1.3.2 Bony fish 32
1.3.3 Amphibians 34
1.3.4 Chicken *(Gallus gallus)* 35
1.3.5 Other birds 40

1.4 Rationale for the analysis of MHC in ducks 43

1.5 Hypothesis and Objectives 45

2 METHODS 46

2.1 Laboratory ducks 46

2.2 Isolation of high molecular weight DNA 46

2.3 Generation of the MHC class I probe (D26E4p) 46

2.4 Generation of the TAP2 probes 49

2.5 cDNA library construction and screening 49

2.6 PCR based sequencing of positive plaques 50

2.7 RT-PCR amplification of MHC class I and TAP2 alleles 51

2.8 Amplification of genomic clones of MHC class I 51

2.9 Amplification of genomic clones of TAP2 52

2.10 Amplification of the intergenic region of MHC class I and TAP2 52

2.11 Southern hybridization 53

2.12 Northern hybridization 53

2.13 Polymerase chain reaction conditions 54

2.14 DNA sequencing and analysis 54

3 RESULTS 56

3.1 Amplification of MHC class I exon 4 sequences from different ducks 56

3.2 Isolation of MHC class I cDNA clones 59
3.2.1 Reverse-transcriptase PCR amplification of MHC class I 64

3.2.2 Northern hybridization 66
3.2.3 MHC class I amino acid sequences 66
3.3 Amplification of MHC class I intron sequences from genomic DNA 72
3.4 Identification of the TAP2 gene 78
3.5 Reverse transcriptase PCR amplification of TAP2 81
3.6 TAP2 amino acid sequences 81
3.7 Amplification and sequencing of the intergenic region between the TAP2 locus and the dominantly expressed MHC class I 86
3.8 Southern hybridization 88
4 DISCUSSION 95
4.1 MHC class I expression in the duck 95
4.2 Polymorphic TAP2 linked to dominant MHC class I locus 97
4.3 The MHC of the duck 99
5 REFERENCES 102
6 APPENDIX A: Nucleotide sequences of MHC class I and TAP2 amplified from genomic DNA of D26 127

List of Tables

Table 1: Oligonucleotides used for the amplification and sequencing of MHC class I and TAP2 products. 47

Table 2: MHC class I alleles amplified from DNA of five individuals from the duck colony 58

Table 3: Pairwise comparisons of the coding region among five duck MHC class I cDNA sequences 65

Table 4: Summary of MHC class I sequences amplified by PCR from genomic DNA. .73

List of Figures

Figure 1: Diagram of the major histocompatibility complex (MHC) class I subunits. 2

Figure 2: Examples of peptides eluted from two mouse MHC class I molecules. 4

Figure 3: Pathway of MHC class I antigen processing and presentation. 6

Figure 4: Organization of the MHC in the chicken (*Gallus gallus*) (not to scale). The region spans approximately 92 kb. Adapted from Kaufman et al. (1999b). 36

Figure 5: Alignment of the MHC class I exon 4 amino acid sequences amplified from different ducks. 57

Figure 6: Nucleotide alignment of the four duck MHC class I sequences from the cDNA library. 63

Figure 7: Northern blot analysis of *Anpl*U expression in different duck tissues. 67

Figure 8: Alignment of the predicted amino acid sequence of the four MHC class I cDNA sequences. 69

Figure 9: Nucleotide alignment of the MHC class I genomic PCR fragments amplified from exon 1 to exon 6. 76

Figure 10: Alignment of duck TAP2 cDNA with chicken TAP2. 80

Figure 11: The predicted amino acid sequence of duck TAP2 alleles A and B. 83

Figure 12: Model of the duck TAP2 protein. 85

Figure 13: Genomic organization of duck TAP2 and the putative major MHC class I locus. 87

Figure 14: MHC class I polymorphism in related and unrelated ducks. 90

Figure 15: Comparison of TAP2 and MHC class I hybridization patterns in five ducks. 93

Figure 16: MHC class I and TAP2 hybridize to restriction fragments of identical size. .94

Abbreviations

ABC	ATP binding cassette
APC	antigen presenting cells
β_2m	beta-2 microglobulin
C	constant
cDNA	complementary DNA
CDR	complementarity determining region
class I-a	classical MHC class I molecules
class I-b	non-classical MHC class I molecules
CTL	cytotoxic T lymphocyte
D	diversity
DHBV	duck hepatitis B virus
ER	endoplasmic reticulum
HBV	hepatitis B virus
H chain	MHC class I heavy chain
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
IL-2	interleukin 2
INF-γ	interferon gamma
ITAMs	immunoreceptor tyrosine-based activation motifs
ITIMs	immunoreceptor tyrosine-based inhibition motifs
J	joining
KIR	killer immunoglobulin like receptor

LMP	low molecular weight protein
MHC	major histocompatibility complex
NBD	N-terminal nucleotide binding domain
NK	natural killer
ORF	open reading frame
PBR	peptide binding region
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RFLP	restriction fragment length polymorphism
Rfp-Y	restriction fragment pattern Y
RT-PCR	reverse transcriptase PCR
TAP	transporter associated with antigen processing
TCR	T-cell receptor
TM	transmembrane
V	variable

1 INTRODUCTION

1.1 Major Histocompatibility Complex (MHC) class I

1.1.1 Structure of MHC class I molecule

The MHC class I molecule is a heterotrimer (Figure 1) composed of the β_2-microglobulin (β_2m) subunit (Peterson et al., 1974), a transmembrane heavy-chain glycoprotein, and an intracellularly derived peptide bound to the heavy chain (Bjorkman et al., 1987a). The classical function of MHC class I molecules is to present peptides to cytotoxic T-lymphocytes (CTL). Recognition of a non-self peptide, in the context of a self-MHC class I molecule, is a signal for CTL mediated killing of the infected cell (Zinkernagel and Doherty, 1974).

Analysis of the crystallized human MHC class I molecules (human leukocyte antigen or HLA) helped define the structure (Bjorkman et al., 1987a; Bjorkman et al., 1987b; Bjorkman et al., 1985; Garrett et al., 1989). For a review of MHC class I structure and function, see Bjorkman and Parham (1990). The heavy chain is composed of three extracellular domains: α1, α2, and α3. The β_2m subunit is bound to the immunoglobulin-like α3 domain of the heavy chain by non-covalent bonds (Peterson et al., 1972; Peterson et al., 1974; Tragardh et al., 1979). The α3 domain provides a contact point for the CTL bound CD8 glycoprotein which is important for cell-cell adhesion between the CTL and target cell (Norment et al., 1988). The α1 and α2 domains are furthest from the cell. They are made up of anti-parallel β-pleated sheets, with long α-helical regions C-terminal to the β-sheets (Bjorkman et al., 1987b). The α-helices form the sides and the β-strands make the floor of the peptide binding region (PBR) (Bjorkman et al., 1987a).

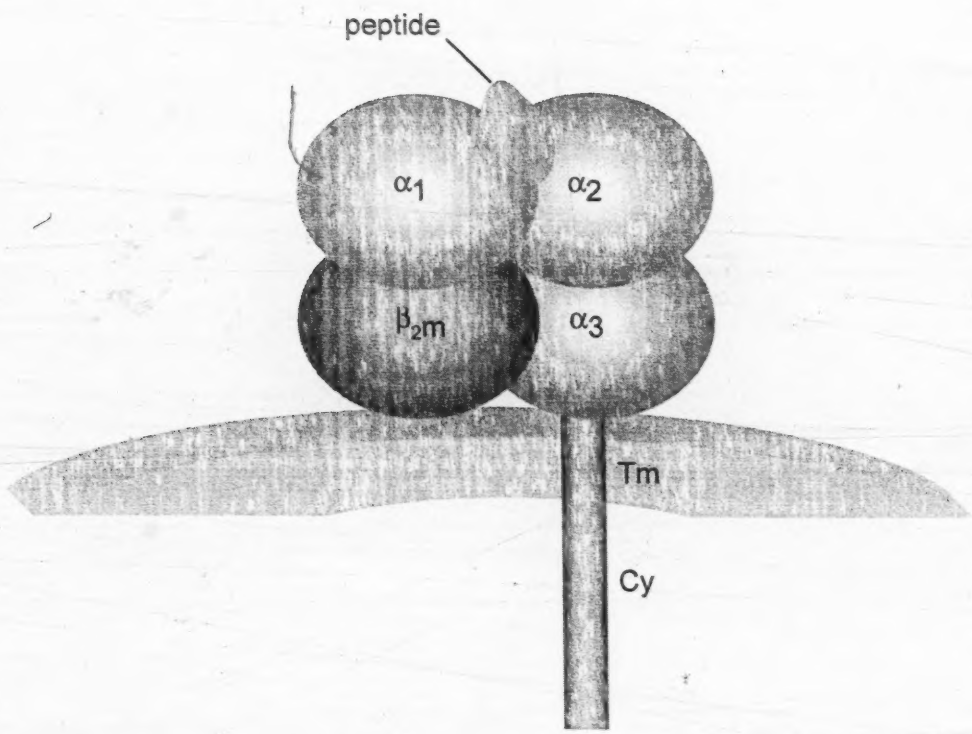


Figure 1: Diagram of the major histocompatibility complex (MHC) class I subunits. MHC class I is a heterotrimer of β_2-microglobulin, a heavy chain transmembrane glycoprotein, and an intracellularly derived and processed peptide. β_2-microglobulin makes contact with the α1, α2, and α3 domains of the glycoprotein, and the peptide rests in a groove formed by the α1 and α2 domains.

Clusters of residues within the narrow ends of the PBR (Saper et al., 1991) are conserved among species (Shum et al., 1999). The structure of peptides and the specificity of the MHC class I PBR were reviewed by Engelhard (1994). The conserved residues of the PBR hydrogen bond with polar main-chain atoms on the carboxy- and amino-terminal ends of the peptide (Bjorkman et al., 1987a; Engelhard, 1994; Garrett et al., 1989; Madden et al., 1991). A peptide must be long enough to span these conserved residues, usually eight residues or greater. However, the closed ends of the PBR also limits the size of peptide bound to around nine residues.

Different MHC class I molecules will only bind peptides that carry the preferred residues in the correct position. Residues within the PBR form pockets that preferentially accommodate certain amino acid side chains from a peptide, based on size and charge. The residues that form the pockets of the PBR are polymorphic between alleles. For each different MHC class I molecule observed, all the bound peptides possess the same residues at invariant positions of the peptide (Figure 2). This gives each MHC class I molecule a characteristic peptide binding motif. The residues of the peptide that need to be present to bind to a particular MHC class I allele are called the anchor residues.

1.1.2 Assembly of MHC class I

Before MHC class I can present peptides to the CTL, its proper assembly is regulated by endoplasmic reticulum (ER) resident proteins. The assembly of MHC class I in humans has been well characterized (reviewed by: Pamer and Cresswell, 1998; Gromme and Neefjes, 2002). The class I heavy chain is highly unstable without a bound peptide and will not be exported to the cell surface without it (see review by Germain and Margulies, 1993). Correct folding of the heavy chain with β_2M and binding of the

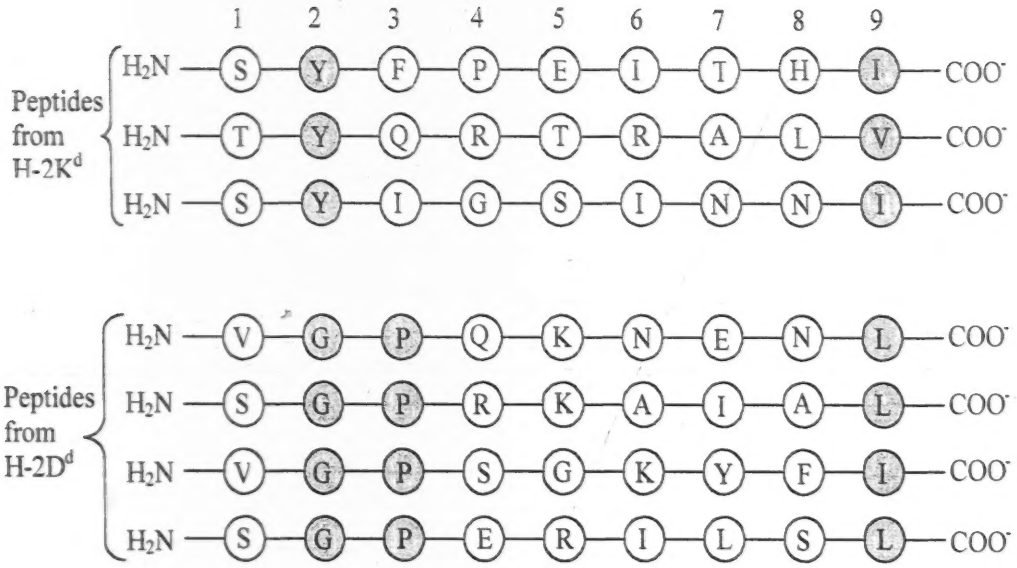

1 2 3 4 5 6 7 8 9
Peptides from H-2K^d
H2N S Y F P E I T H I COO-
H2N T Y Q R T R A L V COO-
H2N S Y I G S I N N I COO-
Peptides from H-2D^d
H2N V G P Q K N E N L COO-
H2N S G P R K A I A L COO-
H2N V G P S G K Y F I COO-
H2N S G P E R I L S L COO-

Figure 2: Examples of peptides eluted from two mouse MHC class I molecules. Shaded residues are the motif residues identified in the sets of peptides eluted. The amino terminus motif residue(s) point into the B-pocket of the MHC class I molecule and the carboxy-terminal residue side chain points into the F-pocket. Adapted from Engelhard, 1994.

peptide to the heavy chain are interdependent events (Boyd et al., 1992; Saper et al., 1991; Townsend et al., 1990; Townsend et al., 1989; Williams et al., 1989). The nascent MHC class I heavy chains (H chain) are translocated into the ER where they associate with the chaperone calnexin (Ahluwalia et al., 1992; Hochstenbach et al., 1992; Ortmann et al., 1994; Rajagopalan and Brenner, 1994) (Figure 3). Calnexin binding facilitates association of β_2m with the H chain, and then dissociates (Ortmann et al., 1994) at which time another ER chaperone, calreticulin, binds to the β_2m/H chain dimer (Sadasivan et al., 1996). Calreticulin appears to remain with the dimer as it binds with the glycoprotein tapasin, which together, facilitate binding with the transporter associated with antigen processing (TAP) (Gao et al., 2002; Grandea et al., 2000; Grandea et al., 1997; Ortmann et al., 1997; Sadasivan et al., 1996; Solheim et al., 1997).

TAP transports peptides from the cytosol into the ER where they can bind to the H chain (Kelly et al., 1992a; Rock et al., 1994). Four tapasin-β_2m/H chain-calreticulin complexes associate with one TAP at a time (Ortmann et al., 1997). Ortmann et al., (1997) hypothesize that this association of multiple MHC class I with TAP may increase the probability that a translocated peptide binds to an H chain. TAP, as a molecule, is not necessary for binding of peptide to the H chain (Henderson et al., 1992; Wei and Cresswell, 1992). However, TAP is needed for transport of antigenic peptides from the cytosol (Kelly et al., 1992a).

Peptides for MHC class I presentation are generated in the cytosol by the digestion of intracellular proteins by the ubiquitin-proteasome pathway (Grant et al., 1995; Rock et al., 1994). During infection, the cytokine interferon γ (INF-γ) stimulates the substitution of subunits LMP2 and LMP7 into the proteasome (Yang et al., 1992).

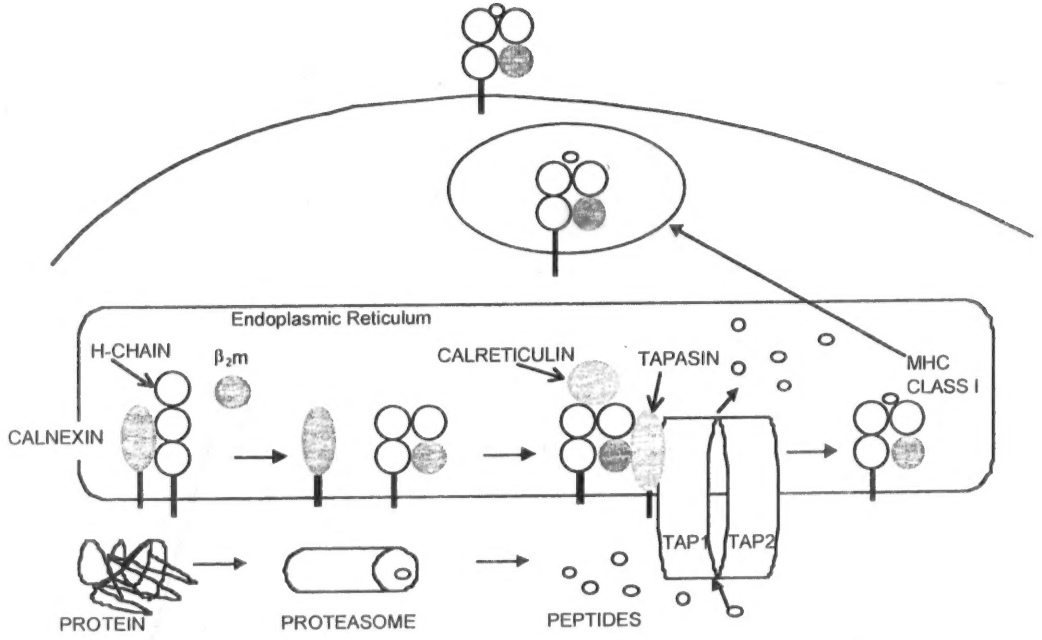


Figure 3: Pathway of MHC class I antigen processing and presentation. Peptides are generated from proteins in the cytoplasm and are transported by TAP into the ER where peptide loading onto the nascent MHC class I heavy chain is facilitated by chaperones. Then the complex is transported to the cell surface where the peptide will be presented to cytotoxic T-lymphocytes in the context of the MHC class I molecule.

Due to this substitution, the number of peptides produced that have basic or hydrophobic carboxy termini (Gaczynska et al., 1994) which MHC class I molecules preferentially bind, is increased (Jardetzky et al., 1991; Madden et al., 1991).

1.1.3 Functions of MHC class I

Classical MHC class I molecules (class I-a) mediate the cellular immune response effected by CTL by binding intracellularly derived antigenic peptides (Bjorkman et al., 1987a; Townsend et al., 1985). However, CTL do not only contact the peptide, they also bind to the MHC. This is critical in determining the population of CTL that are allowed to patrol the periphery. Before a mature CTL is allowed to exit the thymus, it has to be able to bind to MHC through the T-cell receptor (TCR). If the binding is too strong or too weak, the T-cell is deleted from the population by the process of negative or positive selection respectively. In this way, even before they are needed for recognition of potential pathogens, the MHC class I molecules can help select for CTL that will not attack self tissues and cells, and will be able to recognize foreign antigens (reviewed by Benoist and Mathis, 1999).

Other class I molecules, called non-classical or class I-b, do not possess the same level of polymorphism, have a lower level of cell surface expression or may be secreted, and may or may not present antigenic peptides (Shawar et al., 1994). For example, in mice, the receptor FcRn is a receptor on murine neonatal intestine that has affinity for the Fc portion of maternal IgG. FcRn allows the uptake of IgG in the neonate to mediate protective immunity until it establishes its own immune system (Simister and Mostov, 1989). Another example of a murine non-classical MHC class I molecule is H-2M3 which presents N-formylated bacterial peptides instead of viral peptides to the immune

system (Pamer et al., 1992). Non-classical MHC class I molecules can also effect an immune response to pathogens in concert with natural killer (NK) cells or can be factors in the process of graft rejection. These immune functions, as well as non-immune functions of class I molecules, are mentioned in the following sections.

1.1.3.1 *Cell Mediated Immunity*

1.1.3.1.1 T-lymphocytes: T-cell receptor

T-lymphocytes originate from haematopoietic stem cells in the bone marrow stroma (Antica et al., 1994). The T-lymphocyte precursors migrate from the bone marrow to the thymus, where they are called thymocytes. There, they mature either into $CD8^+$ CTL that recognize antigens presented on classical MHC class I molecules or $CD4^+$ helper T-lymphocytes that recognize antigens presented on MHC class II molecules on the surface of antigen-presenting cells (APC) (Benoist and Mathis, 1999). It is the heterodimeric TCR, composed of an α- and β-chain, that makes contact with the MHC molecule and antigen (Garboczi et al., 1996).

The T-cell receptor on the CTL learns to recognize peptides presented in the context of self-MHC class I molecules (Bevan, 1977). The structure and function of TCRs are reviewed by Davis and Chien (1999). Complementarity-determining regions (CDRs) 1 and 2 are the primary contact points with the α1 and α2 domains of the MHC class I molecule and the conserved terminal anchor residues of the peptide (Garboczi et al., 1996). The highly variable CDR3 makes contact with the peptide presented on MHC class I (Garboczi et al., 1996). The variability in the CDR3 allows different TCRs that recognize the same MHC molecule, to recognize the array of peptides that could be bound to that MHC molecule.

The α- and β-chains of the TCR are immunoglobulin-like in structure (Novotny et al., 1986) and genomic organization (Gascoigne et al., 1984; Hayday et al., 1985). Four regions contribute to the structure of the β-chain: the V (variable), D (diversity), J (joining) and C (constant) regions (Gascoigne et al., 1984). The α-chain does not contain a D-region (Hayday et al., 1985). For each region of a chain, there are a variety of gene segments within the genome; very few D (none for the α-chain) and C segments but many V and J segments (Davis and Bjorkman, 1988).

The highly variable CDR3 that binds the peptides presented on MHC class I, is generated by the V(D)J joining. Each T-cell expresses a TCR that is generated by random rearrangement and joining of segments from each of these regions (Akira et al., 1987). For each chain of a TCR that is expressed by a T-lymphocyte, each region is randomly chosen from the variety of gene segments present in the genome (reviewed by Davis and Bjorkman, 1988). The diversity generated by randomly linking different segments from each region together is augmented by random addition of nucleotides at each junction called N-nucleotide addition (Alt and Baltimore, 1982). Also, rearrangements in the β-chains can result in one or two D segments being added (Kavaler et al., 1984). There is a limited amount of diversity contributed by the small number of D segments available for the β-chain. This is compensated for by coding of the D segment in all three reading frames. The number of possible gene segment combinations for each α- and a β-chain that make up the TCR, means that there could be around 10^{15} αβ TCR combinations possible (Davis and Bjorkman, 1988). Selection against T-cells expressiong TCRs that bind too strongly to self MHC or self peptide, means that many of the possible TCR combinations will not be expressed on mature T-cells.

A third type of T-lymphocyte that matures in the thymus expresses a TCR made up of two different chains: γ and δ (reviewed by Davis and Chien, 1999). These lymphocytes were discovered after the cloning of the δ-chain genes (Chien et al., 1987a) located 5' of the $J_\alpha C_\alpha$ genes. The γδ lymphocytes appear before the αβ (Chien et al., 1987b; Pardoll et al., 1987), and are mostly associated with epithelial cells of the skin (Koning et al., 1987), intestine (Goodman and Lefrancois, 1988), female reproductive tract (Itohara et al., 1990) and lung (Augustin et al., 1989). It is known that γδ lymphocytes are more abundant in birds than mammals (Bucy et al., 1991). The function of γδ lymphocytes is different from αβ lymphocytes because they do not respond to peptide antigen presentation, but respond to the level of the MHC expression on a cell (Schild et al., 1994) or the expression of a ligand on a stressed cell (Bukowski et al., 1994).

1.1.3.1.2 T-lymphocytes: Development

The β-chain of the TCR is first expressed on the surface of the immature T-cell, linked through a disulfide bond to the protein gp33 (Groettrup et al., 1993). It is thought that the binding of this pre-TCR molecule to an unidentified intrathymic ligand can inhibit further β-chain rearrangements and induce the expression of the co-receptors CD4 and CD8 (von Boehmer, 1994). Also, the expression of the pre-TCR may enhance the rearrangement of the α-chain (Groettrup et al., 1993).

With the αβ TCR expressed on the thymocyte surface, further maturation of the thymocyte is provoked. Positive selection selects for T-cells that can recognize self-MHC molecules (Matzinger et al., 1984; Zinkernagel et al., 1978). The TCR must be able to bind to an MHC class I or II molecule or it is deleted from the thymus. However,

if the binding to a self-MHC molecule, or the MHC-bound self-antigen, stimulates the T-cell to activation, negative selection will cause the T-cell to be deleted (Ignatowicz et al., 1996; Kisielow et al., 1988). It is believed that positive and negative selection in the thymus is under control of the avidity (the number of interactions between a TCR and self-MHC) and affinity (the strength of the reaction) (Janeway et al., 1992). That is, if there are a low number of high affinity TCR-MHC-peptide complexes formed, the T-cell will not be selected against. However, a high number of these interactions results in cell death.

The selection of TCR by MHC in the thymus is what limits the number of MHC loci expressed by any organism (Lawlor et al., 1990). If MHC molecules only presented peptides to be surveyed by the immune system, one would assume the greater the number of different MHC molecules expressed, the greater chance of eliminating pathogens. However, the function of eliminating T-cells that react and kill cells expressing self, means that if there are more MHC molecules, more T-cells will be deleted in the thymus, decreasing the overall effective T-cell population available.

T-cells that do make it to the periphery may still be selected against. They need to encounter antigen with a specific MHC molecule in order to proliferate, but studies have shown that constant stimulus with the same antigen can cause a type of negative selection called clonal anergy (Rocha and von Boehmer, 1991). This can help an organism in cases where a T-cell has an affinity to self molecules but is not eliminated and escapes the thymus. However, if an organism is chronically infected by a pathogen, the chronic presence of the pathogen in the periphery may also cause the T-cells that recognize it to become anergic.

1.1.3.1.3 Graft rejection

The rejection of transplanted organs between two individuals of the same species is called allograft rejection. Research in this area is quite developed because of the medical relevance for organ transplantation. The cell mediated immune system of a vertebrate is not able to recognize that an organ transplant was done out of necessity, only that it is foreign. However, the alloreactive response that occurs is much stronger than an immune response against a pathogen invader. For a review of this topic, see Auchincloss et al., (1999).

Exposure of a T-cell to an alloantigen can occur by the presentation of a donor antigen on donor MHC. T-cells can also be exposed to antigens from a graft by presentation of the donor antigens on self-MHC. The allogenic peptides presented can be derived from the allo-MHC (Gallon et al., 1995). They can also be minor histocompatibility antigens (Wallny and Rammensee, 1990) derived from any protein that possesses allelic variants between individuals. This can include non-classical MHC class I molecules.

Pregnancy may not be considered a graft in the traditional sense, but it is a situation where an individual carries tissue that is not self. In this case, it is also evolutionarily inadvisable to try to get rid of every foetus that differs at the MHC, since it is sexual reproduction that encourages differences. Unlike transplantation, a foetus is protected from maternal immune system attack, in at least one way, by the maternal MHC. The non-classical class I gene product HLA-G is expressed on the extravillous trophoblast surrounding the foetus. HLA-G inhibits NK cell activity, which would

otherwise detect the absence of the maternal-MHC on the foetus and target it as foreign (Yokoyama, 1999).

1.1.3.1.4 Natural Killer cells

Natural killer cells originate from a lymphoid precursor in the bone marrow that is also thought to give rise to T cells (Spits et al., 1995). These cells were first recognized by their cytotoxic activity towards tumours, and are now known to also attack virally infected cells that have altered expression of MHC class I (Storkus et al., 1989). Viruses are able to alter the expression of MHC class I on the surface of cells (Wiertz et al., 1997) and thus inhibit activation of cytotoxic T-lymphocytes. However, NK cell activity can be inhibited or activated by recognition, or lack of recognition, of different classical or non-classical MHC class I molecules and their peptides on the surface of the cell. The interaction between NK cell receptors and MHC class I or stress induced ligands is reviewed by Lanier (1998), and Biassoni and colleagues (2001).

In humans, receptors from the killer Ig-like receptor (KIR) family and the leukocyte Ig-like receptor (LIR) family are inhibitory or activatory receptors that recognize specific HLA alleles (see review by Biassoni et al., 2001). In mice, the C-type lectin-like Ly-49 receptor family are dependent on classical MHC class I expression in order to mediate NK cell killing (Brennan et al., 1996).

Inhibitory NK cell receptors contain immunoreceptor tyrosine-based inhibition motifs (ITIMs) on the intracellular portion of the receptor. When the receptors are cross-linked by the MHC class I ligands, the ITIM is stimulated to send a signal to the cell that inhibits cytotoxic activity (Lazetic et al., 1996). Some NK cell receptors lack the ITIM motifs, and have been implicated in activating NK cell mediated killing. These receptors

are associated with either the DAP10 or DAP12 intracellular adaptor proteins. DAP10 and DAP12 possess immunoreceptor tyrosine-based activation motifs (ITAMs) that stimulate different intracellular pathways leading to NK cell activation (Wu et al., 2000).

An inhibitory and an activating receptor can have affinity for the same ligand. For example, the class I-b molecules HLA-E or Qa-1 in humans or in mice respectively, bind signal sequence peptides cleaved from expressed MHC class I molecules (Borrego et al., 1998; Braud et al., 1998; Lohwasser et al., 2001). The inhibitory receptor heterodimer CD94/NKG2A and the activating receptor CD94/NKG2C combined with DAP12, both recognize these non-classical MHC class I molecules in humans and mice. The presence of an activating receptor that binds the same ligand as the inhibitory receptor appears to contradict the inhibition. It is possible, since all NK cells possess at least one inhibitory receptor, stimulation through the inhibitory receptor over-rides activation of the cell by the activating receptor. For example, viruses may subvert antigen presentation by MHC class I by down-regulating its expression. The inhibitory receptor would not be stimulated by a large number of ligands. However, a small number of ligands still present could stimulate the activating receptor, allowing NK cell activation (Lanier et al., 1998).

Another example of NK cell recognition of a non-classical MHC class I molecule is the receptor NKG2D. This receptor in humans and mice recognizes expression of the class I-b molecules MICA and MICB (Bauer et al., 1999; Wu et al., 1999) or Rae-1 and H60 (Cerwenka et al., 2000) respectively. Expression of the NKG2D ligands are induced by stress and are often found on virus-infected cells or on tumours. Cross-linking of the

NKG2D receptor mediates killing of the stressed or transformed cells through stimulation of intracellular adaptor protein DAP10.

1.1.3.2 Non-immune functions of MHC

The immune responses of MHC class I and other MHC encoded molecules are well known. However, the MHC also seems to influence functions not related to the immune system. Lack of MHC class I molecule expression reduces insulin uptake by cells (Due et al., 1986). In human (Kittur et al., 1987) and mouse cells (Verland et al., 1989), insulin receptor binding affinity has been directly correlated with the expression of different MHC class I alleles.

Experiments have also shown that cell-to-cell contact can be mediated by MHC class I expression. When monolayers of mouse epithelial cells are grown, if two groups of cells with different MHC class I haplotypes are grown next to each other, there is a significant decrease in the number of overlapping cells compared to groups of cells with identical MHC (Curtis and Rooney, 1979).

Other ways that the MHC can influence an organism may be caused indirectly by an effort to maintain genetic diversity, especially at MHC loci, in a population. The following sections touch on how the MHC influences an individual to choose a suitable mate, or decide who are kin. As well, post-fertilization methods that may select for MHC diversity (or against MHC homogeneity), and how MHC influenced sexual selection varies between species will be dealt with.

1.1.4 Diversity in MHC class I

Different MHC class I alleles can have different peptide binding motifs. This means that different molecules expressed on the surface of the cell can present a different

sampling of peptides to the CTL. Pockets along the groove bind the side chains of peptides (Saper et al., 1991) eight to eleven amino acids in length (Falk et al., 1991). Pockets within the cleft, especially ones that bind peptide residues near the termini, show specificity for certain amino acid side chains (Falk et al., 1991). Other pockets within a molecule are able to accommodate a variety of amino acid side chains (Jardetzky et al., 1991), which gives the molecule the ability to bind a large variety of different peptides. Between different MHC class I molecules, there is a greater number of variable residues in positions that contact peptide, than in other parts of the molecule (Bjorkman et al., 1987a; Parham et al., 1988). In this way, the motifs of the bound peptide differs from allele to allele, (Falk et al., 1991; Garrett et al., 1989; Parham et al., 1988).

To combat the need to keep the number of MHC class I genes low (so a large T-cell population can be selected for) and still be able to effectively identify pathogens, a population can exhibit great allelic diversity. For the HLA-A, B, and C loci in humans, there have been 266, 511, and 128 alleles identified at each locus respectively (Robinson et al., 2000). The nucleotide sequences of alleles from the same locus are more similar to each other than alleles from other loci (Lawlor et al., 1988; Parham et al., 1989). The 3'-untranslated region, the cytoplasmic domain, and the transmembrane domain exhibit locus specific traits (Parham et al., 1989). This is not surprising since there are fewer non-synonymous nucleotide substitutions in these regions compared with the domains that form the peptide-binding cleft (Hughes and Nei, 1988). The following sections discuss how this allelic polymorphism may be selected for.

1.1.4.1 Molecular mechanisms of generating MHC diversity

Between MHC class I alleles, extreme polymorphism in the peptide binding regions of MHC class I occurs because of selection. This is determined by the fact that the number of non-synonymous amino acid substitutions exceeds the number of synonymous substitutions (Hughes and Nei, 1988). However, the level of neutral mutations in the MHC does not exceed that of the rest of the genome (Parham and Ohta, 1996).

There are two possible theories for the evolution of MHC polymorphism. The first is based on the idea that new species arise from small populations of old species (Lawlor et al., 1990). In this instance, most pre-existing MHC polymorphism is denied to the new species and they must develop new polymorphism. If this is the only mechanism by which polymorphisms are generated, the majority of MHC polymorphism seen today would have to have been generated by mutation.

The second view allows for accumulation of MHC polymorphism by a normal rate of point mutation (Ayala, 1995; Parham and Ohta, 1996). This theory is based on trans-species (trans-specific) evolution (Arden and Klein, 1982; Ayala, 1995), where polymorphisms that existed in the ancestral species are transferred to the new species. An example of this is the lack of divergence between MHC class I alleles of humans and chimpanzees (Lawlor et al., 1988). However, side by side, chimp and human MHC class I alleles are not distinguishable as being from different species, revealing a common ancestral origin. The combination of trans-species evolution with other mechanisms of generating MHC polymorphism, such as point mutation and gene conversion, over time

can increase significantly the numbers of alleles in a population, and increase their divergence from each other (Parham et al., 1988).

In humans, the most common form of gene conversion appears to be interallelic conversion (Parham et al., 1988). This has occurred when two alleles differ in a short stretch of nucleotides, and the stretch from one of these alleles is identical to a third allele (Parham and Ohta, 1996). This mechanism may be responsible for the homogenization of non-coding parts of alleles at a locus and the generation of locus specific traits (Parham et al., 1989). It may also contribute to diversity through recombination of polymorphic regions between alleles.

MHC class I loci in mice do not show locus specific traits to the same extent as humans, and alleles from different loci can appear more similar than alleles from the same locus (Lawlor et al., 1990). Inter-locus gene conversion, identified by the sharing of motifs between alleles from different loci, is thought to be responsible for this homogenization of the MHC class I alleles from different loci in mice (Kuhner et al., 1990). This type of gene conversion is also able to maintain allelic diversity by re-assorting the variability present among loci, including those at non-classical class I loci (Watkins et al., 1990) and non-functional loci (pseudogenes).

1.1.4.2 Natural selection for MHC polymorphism

Pathogens can drive selection for MHC polymorphism because MHC-heterozygotes (heterozygote advantage) (Doherty and Zinkernagel, 1975) and/or rare alleles (Bodmer, 1972; Slade and McCallum, 1992) have a better chance of survival against a pathogen. If the MHC allele carried by an individual is not useful in eliminating a specific pathogen, a heterozygote at that locus could have an advantage

over a homozygote. The selection of a heterozygote individual over a homozygote is called over-dominant selection. An example of the advantage of MHC heterozygosity is clear in humans infected with human immunodeficiency virus (HIV). MHC heterozygotes infected with HIV have a longer lifespan than MHC homozygotes (Carrington et al., 1999). MHC heterozygosity increases the chances that an individual can recognize and eliminate a pathogen before the HIV weakened immune system is overtaken by it.

The susceptibility of an individual that carries an ineffective allele can be masked by selection for the resistance trait the allele might possess. This is called frequency-dependent selection. The ineffective allele also might be masked by the phenotype of alleles at other loci, thus, allowing the ineffective allele to remain in the population (Potts and Slev, 1995).

In mice, entrenched mechanisms of avoiding inbreeding depression such as familial imprinting (Manning et al., 1992; Penn and Potts, 1998; Yamazaki et al., 1988) can lead to selection for MHC heterozygotes. We also see in humans, mating preferences (Ober et al., 1997) may help maintain MHC diversity in a population. Avoidance of inbreeding, to prevent depression of genetic diversity, and the selection of heterozygosity, as a mechanism for decreasing parasite/pathogen load, are methods of sexual selection that can be explained by the Hamilton-Zuk model (Hamilton and Zuk, 1982). The Hamilton-Zuk model states that individuals choose mates based on their disease resistance, which is reflected in the potential mates' physical characteristics. A recent review (von Schantz et al., 1999) connects the condition dependent traits in male birds (such as coloured feathers and keratinised cells of a spur), to the oxidative stress on the

individual, when infected by a pathogen. Oxidative mechanisms are evoked by innate immunity as a first response to a pathogen. If the individual does not have an adequate cell-mediated immune response, effected by the MHC, then the continuing oxidative stress will affect the individuals' appearance. They suggest that mate choice is based on these characters to select for effective MHC. There are some studies that support this idea. For example, studies showing pheasant spur length can be correlated with MHC genotype and subsequent female mate choice (von Schantz et al., 1996) which will be discussed later. Also, female sticklebacks that mated with more brightly coloured males, produced offspring more resistant to parasites (Barber et al., 2001).

The other mechanism of generating MHC diversity by pathogen selection is frequency-dependent selection. Note that this model is not in conflict with the heterozygote advantage model, but can work in concert. In the frequency-dependent selection model, a rare allele in a population can become dominant because it offers resistance to a particular pathogen (Slade and McCallum, 1992). As the allele becomes more frequent, the pathogen can develop resistance to it and thus select against it. The cycle may begin once more when the allele is rare again and is no longer selected against by the pathogen (Potts and Slev, 1995). These types of events are apparently intrinsic to host-parasite interactions (Slade and McCallum, 1992). An example of this in humans could be the elevated frequency in The Gambia of MHC class I and II alleles associated with resistance to malaria (Gilbert et al., 1998; Hill et al., 1991).

1.1.4.2.1 Sexual selection

1.1.4.2.1.1 MHC class I related odours

The MHC of an individual is believed to influence mate choice by MHC-linked odour. MHC antigens are detectable in the urine of mice and rats (Singh et al., 1987). Rats familiarized to the odour of urine from individuals with a certain MHC class I allele, showed reactions expressing unfamiliarity when exposed to the odour of congenic rats, genetically identical to the first group except for carrying a different MHC class I allele (Singh et al., 1987). Mice, when given a choice, spend more time near the urine odour from individuals that carried MHC alleles different from that of their maternal genetic strain (Isles et al., 2001). MHC class I haplotypes can be correlated to the differing ratios of volatile chemicals detected electronically in the urine (Montag et al., 2001). Even when urine is heated, denaturing the MHC molecules, mice still show the ability to choose between different congenic strains (Singer et al., 1993). It is suggested that MHC molecules could bind allele specific volatile molecules that could be excreted in the urine (Singer et al., 1997; Sobst et al., 1999). Other research has also detected soluble MHC molecules in human serum, sweat, saliva, and tears (Aultman et al., 1999; Sobst et al., 1999).

Genomic data shows linkage between a group of polymorphic olfactory receptor genes and the MHC class I region in humans and mice (Ehlers et al., 2000; Younger et al., 2001). The allelic polymorphism of the olfactory receptor genes could segregate with the MHC loci, possibly leading to functional linkage of these genes. Other research (Loconto et al., 2003) shows a family of non-classical MHC class I molecules in mice, function as escort molecules for a family of pheromone receptors. Male mice lacking

these non-classical MHC class I molecules did not show characteristic aggressive behaviour towards other males. This body of data builds the case for the ability of vertebrates to detect MHC differences from body odour.

1.1.4.2.1.2 Familial MHC imprinting

The MHC disassortative mating hypothesis proposes that a pair will avoid mating with each other if MHC similarities are detected. Some studies that favour the odour-linked MHC disassortative mating hypothesis, show male mice, when given a choice, choose mates that do not share the same MHC as the mother that reared them (Manning et al., 1992; Yamazaki et al., 1988). A study of semi-natural mice (Penn and Potts, 1998) showed that females fostered in an environment with MHC dissimilar foster-family members, were more likely to mate with MHC similar mates than expected. This also confirms the results of another study showing mice that were not fostered, (they stayed with their mother) exhibited high levels of MHC-disassortative mating (Potts et al., 1991).

It is suggested that familial imprinting of MHC associated odours is a mechanism to increase heterozygosity in offspring, as well as avoid inbreeding. Human mate choice also seems to be governed by this. Mate choice in Hutterites, a population of reproductively isolated humans that originated from a small founder population, showed individuals chose mates dissimilar at MHC class I and II loci (Ober et al., 1997). Among couples that did match for a MHC haplotype, the common haplotype in the female was less often inherited from her mother than her father. This suggests that maternal imprinting in humans could be as important in mate choices as it is in mice (Manning et al., 1992; Penn and Potts, 1998; Yamazaki et al., 1988). However, another study on

South Amerindian tribes found that mating in the tribe was random with no HLA preference (Hedrick and Black, 1997). The controversy surrounding the assumptions of these studies, is that in human society, especially in isolated populations, there can be confounding cultural factors influencing mating practices (Beauchamp and Yamazaki, 1997).

Some interesting studies in humans tested preferences for the odour of MHC similar or dissimilar individuals by asking testers to smell worn T-shirts. The results of the first studies were similar to mice, suggesting that human mate selection could be dependent on MHC differences, which are detected by preference to certain body odours (Wedekind and Furi, 1997; Wedekind et al., 1995). They showed the preferences of males and females for an odour decreased with an increase in the number of MHC alleles shared.

Females may prefer mates slightly similar at the MHC, selectively maintaining some of the alleles they have inherited, or preventing the break-up of genes that may have co-evolved as a strategy against pathogens (Ochoa and Jaffe, 1999). A study using the Hutterite population previously mentioned (Jacob et al., 2002), measured individuals' preferences for the odours of worn T-shirts, similar to the studies by Wedekind et al. (1995, 1997). They showed odours of individuals were disliked if the smeller shared less than one, or greater than two alleles with the donor. When a female preferred the odour of a male carrying one or two HLA alleles that matched their own, the alleles that were common to the female recipient and the male donor were ones the females had inherited paternally (Jacob et al., 2002). This seems contrary to the previous studies (Wedekind and Furi, 1997; Wedekind et al., 1995) that suggested preferred odours are of potential

mates that are not similar at the MHC. However, the studies by Wedekind et al. did not determine the complete MHC haplotypes of the individuals tested, and they did not track maternal and paternal inheritance of alleles. An earlier study of mating preference in mice suggested that female mice prefer to mate with males that had some similar genes or were not completely different (Gilder and Slater, 1978). For humans, even though odour is likely very important in choosing a desirable mate there are other possible decision mechanisms in place.

1.1.4.2.1.3 Maternal-foetal interactions

There can still be selection for MHC dissimilarity after mate selection. A study of some South Amerindian tribes, with apparently random mating, showed there were more heterozygote offspring than expected under normal Mendelian expectations (Black and Hedrick, 1997). The selection against homozygote offspring occurred when the parents shared a haplotype, but the mother was heterozygous and the father was homozygous. The selection did not occur if the mother was homozygous and the father was heterozygous. Results from a Hutterite population also revealed fewer HLA homozygote offspring than expected (Kostyu et al., 1993). Infertility studies (Ho et al., 1994) report that couples undergoing infertility treatment more often lose a foetus if they share more than one HLA allele. However, the mechanism regulating the maternal-foetal interaction that selects against HLA homozygotes is still unknown.

1.1.4.2.1.4 Sexual selection in non-mammalian vertebrates

In non-mammalian vertebrates, research concerning the connection of sexual selection with the MHC is not as extensive but shows similar features to mammalian mate preferences. A study of a large number of pheasants showed that male spur length,

a secondary sexual characteristic, was significantly correlated with certain MHC haplotypes (von Schantz et al., 1996). The same group had previously reported that female pheasant mate choice is also correlated with male spur length. Von Schantz et al. (1996) suggest that by choosing a mate with long spurs, a female increases the likelyhood of survival of her offspring. The appearance and attractiveness of the secondary sexual characters on a bird, like a spur, is influenced by the health of the bird (von Schantz et al., 1999). If parasite or pathogen load is high, the individual is not able to divert as much energy to the appearance or maintenance of the characters. In this way, it is hypothesized that these characters, which can be predictive of male viability, are linked to MHC genotype. However, the MHC linkage of these characteristics in birds is more tenuous then in other vertebrates.

Sticklebacks are believed to be able to use odour to choose mates with a larger number of MHC class II loci (Reusch et al., 2001). However, this study may be flawed because the measurement of the number of loci carried was based on genomic sequences, not expressed sequences. Female fish were scored as preferring a male based on the amount of time she spent near the water that flowed from the tank the male had previously been in. The female stickleback showed they were interested in the odour of the potential mates by releasing eggs after exposure to water males had been in previously. The number of alleles an individual had was determined by amplifying the peptide binding region of MHC class II loci with common primers.

It is possible that fish can choose mates, preferring those with a variety of alleles, that is, choosing a heterozygous mate over a homozygote. A study of Atlantic salmon typed the alleles at five microsatellite loci and an MHC class II locus in a group of

spawning salmon, and then analyzed the genotypes of the offspring to determine the parental genotypes (Landry et al., 2001). They showed a preference for mates with MHC alleles that differed from their own at the peptide binding region, but with no specific selection against inbreeding at the level of the genome.

Studies relating sexual selection with MHC in a variety of vertebrates will likely become frequent as confidence in the mammalian models increase, along with extensive characterization of the MHC in many species. However, it is important to remember that mate selection is not just based on one trait. For example, chemical traces in urine from molecules of non-MHC related genes, such as those determining sex, can contribute confounding variables when studying urine components (Montag et al., 2001). Also, studying human mate choices based on MHC can be confounded by our social structure and habits (Beauchamp and Yamazaki, 1997). Just as mate selection is not based only on MHC diversity, MHC diversity is not based solely on sexual selection.

1.2 Transporter associated with antigen processing 2 (TAP2)

1.2.1 Structure and function

The antigen-processing molecule, TAP, is a heterodimer of TAP1 and TAP2 (Kelly et al., 1992b) within the ER membrane (Kleijmeer et al., 1992) that transports antigenic peptides from the cytosol into the ER (Spies et al., 1992). It is related to a family of transporters characterized by possession of the highly conserved ATP-binding cassette (ABC) (Trowsdale et al., 1990) and requires the binding and hydrolysis of ATP to be active (Neefjes et al., 1993). Each subunit possesses an N-terminal nucleotide binding domain (NBD) with the characteristic ABC sequence motifs, however, the

binding and hydrolysis of ATP on TAP2, not TAP1, is believed to be the factor that limits peptide binding to TAP (Karttunen et al., 2001; Saveanu et al., 2001).

The ABC is identified by highly conserved sequence motifs (Schneider and Hunke, 1998). The Walker A (GxxGxGKT/S) and B (hhhhD, where h stands for hydrophobic amino acids) sites are responsible for the binding and hydrolysis of ATP (Walker et al., 1982). Another sequence, unique to this family of transporters called the signature sequence or linker peptide (LSAGQ(Q/K/R)QR), is situated between the Walker A and B sites (Hyde et al., 1990; Prosite). It is thought that the signature sequence on one half of a transporter interacts with the Walker A sequence on the opposing half (Loo et al., 2002). The displacement of this sequence upon hydrolysis of the nucleotide could cause the conformational change of the ABC transporter (Loo et al., 2002), and in the case of TAP, lead to translocation of the peptide into the ER.

The molecular structure of ABC transporters, a family of membrane-spanning proteins, is not well characterized. The topology of TAP1 and TAP2, determined by mapping the regions of hydrophobic amino acids that could possibly constitute the transmembrane (TM) segments, yields results very different from truncation analysis. If an even number of TM segments were predicted, by default the N-terminus would be in the cytoplasm because the C-terminus NBD is in the cytoplasm. Groups using hydropathy analysis (Momburg et al., 1996; Ohta et al., 1999) determine TAP1 and TAP2 have ten TM segments each. Analyzing truncated molecules, another group (Nijenhuis and Hammerling, 1996) reported both molecules to possess six TM segments.

Recently, imaging analysis of the TAP heterodimer predicted each half of the transporter to be large enough to accommodate seven or eight TM segments (Velarde et

al., 2001). Other analyses of truncated molecules predicted that the N-terminus of TAP2 is in the ER lumen, and proposed that TAP1 has eight TM segments while TAP2 has seven (Vos et al., 1999) which correlates well with the predicted topology from imaging analysis. The same group has also proposed a head-head/tail-tail orientation of the TAP subunits (Vos et al., 2000). They have shown the six C-terminal TM segments of TAP1 and the five C-terminal TM segments of TAP2 interact to form the pore of TAP, with the NBDs of each subunit on the same side of the pore.

The genes of TAP1 and TAP2 were originally found in the human MHC class II region (Trowsdale et al., 1990) and are located in the MHC class II region of many mammals, far from the MHC class I antigen presenting molecules. The rat is an interesting exception to this. The expressed MHC class I loci of the rat are within the class II region, linked to the antigen processing genes like TAP1 and TAP2 (Gunther and Walter, 2001; Powis et al., 1992). It is possible in the rat, that the MHC class I and TAP genes, which are linked physically and functionally, have co-evolved together (Joly and Butcher, 1998). Studies have showed that some MHC class I and TAP2 alleles have preferences for peptides with differently charged C-termini.

The rat MHC class I allele $RT1^a$ (as well as alleles d, f, g, l, q, s), prefers to bind peptides possessing a C-terminal arginine whereas $RT1^b$ (also c, h, k, m, n, u) prefers to bind peptides with only a hydrophobic C-terminal residue (Joly et al., 1994). Since peptide transport into the ER by TAP is known to be dependent on the charge of the C-terminus (Neefjes et al., 1993), the specificity of TAP can affect whether the MHC class I molecules can find suitable peptides to bind or not. Rat TAP2A (refers to alleles $TAP2^a$ and $TAP2^l$) has been shown to bind and transport peptides with hydrophobic or basic C-

termini, whereas TAP2B (alleles TAP2^c and TAP2^u) prefers peptides with only hydrophobic C-termini (Momburg et al., 1994). When TAP2^b alone is linked to RT1^a, the class I molecule does not bind peptides efficiently and remains in the ER (Powis et al., 1991).

There are 25 amino acid differences between the two different rat TAP2 molecules but only a few are involved in the differential binding of peptides. The peptide binding site is thought to fall on the cytoplasmic loop between the two TM segments closest to the NBD and after the last TM segment in the cytosol (Nijenhuis and Hammerling, 1996). The amino acid changes A217T and E218M between RT1^a and RT1^b have been implicated by separate groups as causing the differential transport of peptides (Deverson et al., 1998; Momburg et al., 1996), however, both groups also point to other positions that are important. Momburg *et al* (1996) showed that S374E and Q380R contribute to the peptide binding differences of the two TAP2 alleles, but Deverson et al. (1998), could not find evidence for the two alleles. However, they did show that the amino acid changes Q262R, S265P, and L266F contribute to transport differences. Although it is clear that TAP2 linkage with MHC class I can alter the ability of an organism to effectively detect pathogens, further work needs to be done to clearly define the structure and differential function of the transporter in other vertebrates.

1.3 Organization of the MHC in vertebrates and functional consequences

Adaptive immunity is not limited to mammalian vertebrates, although much of the research attempting to elucidate the function and associations of molecules encoded in the MHC, focus on well-known mammalian systems: human cell lines, mice, and rats. The human MHC is, so far, the best-characterized MHC region among vertebrates. The

completion of the genomic map of the region shows it is very complex, with the identity of many genes remaining unknown (MHC Sequencing Consortium, 1999). Class I, class III and class II genes are located in well-defined regions telomeric to centromeric respectively on chromosome 6, and span approximately 3.6 Mb. However, the organization and diversity of genes in these regions in humans is not representative of all other vertebrates. This topic is reviewed by Flajnik and Kasahara, 2001, and Flajnik et al., 1999.

The difference in MHC organization among vertebrates can be illustrated by examining the organization of the antigen processing gene, TAP2, and the MHC class I genes. In humans, TAP genes are located in the class II gene region, separate from MHC class I (Kelly et al., 1992a). Compare this organization with the rat, that has a cluster of class I genes situated in the class II region, near the TAP loci (reviewed by Gunther and Walter, 2001) which has allowed the class I and TAP genes to co-evolve (Powis et al., 1996). With the genes separated in humans, the MHC class I genes could develop a wide range of antigenic peptide binding specificities, as long as TAP was permissive for a variety of peptides. The observation that TAP and MHC class I genes are linked in many non-mammalian vertebrates and in some mammals, suggests that the organization in humans is a fairly recent phenomenon.

Available data exploring the evolution and adaptation of the MHC does not represent the evolutionary history of different vertebrate lineages. Often subjects are chosen based on economic importance or need for conservation. With growing interest in this area, studies on species more representative of different lineages, may help us better

understand how the adaptive immune system of an organism functions under pathogen pressure.

1.3.1 Cartilaginous fish

Sharks, a cartilaginous fish, are the oldest living jawed vertebrate (gnathosome) (Kumar and Hedges, 1998), and the oldest organism that possesses MHC class I antigen processing and presentation genes. Class I and class I-like genes have been identified in different species of shark. In the horned shark (*Heterodontus francisci*), an expressed classical MHC class I gene and a non-classical class I gene were identified. The non-classical class I gene appears to be closely related to a non-classical class I gene found in the nurse shark (*Ginglymostoma cirratum*) (Bartl et al., 1997). In the banded houndshark (*Triakis scyllia*) (Okamura et al., 1997) at least one class I locus was found at which alleles possessed the necessary motifs to be classical antigen presenting molecules and in some individuals a second locus was present. Most recently in the spiny dogfish shark (*Squalus acanthias*), a highly divergent non-classical class I gene was identified (Wang et al., 2003). The extreme divergence of this gene from known shark class I sequences have caused some to postulate that this sequence shows sharks carry non-classical alleles with very diverse immune functions, such as the human CD1 or mouse MICA (Wang et al., 2003).

The the region encoding the class I molecules in sharks has not been sequenced completely, but a number of MHC genes are known to be linked. Classical class I, class II α and β (Ohta et al., 2000), proteasome and TAP genes (Ohta et al., 2002) are known to be closely associated, with non-classical class I loci located outside of this region

(Ohta et al., 2002). The linkage of these genes in many non-mammalian and mammalian vertebrates suggests that these genes have been functionally linked through evolution.

1.3.2 Bony fish

A group of bony fish in which the MHC has been studied is the Teleostei. This extremely diverse group diverged from the lineage that led to the Coelacanthiformes and tetrapods, approximately 450 million years ago (Kumar and Hedges, 1998). The MHC has been studied in species from a number of different Teleost orders. Even though over a hundred million years can separate some of these species from their most recent common ancestor, there are features of the MHC that are conserved among them.

In terms of organization, the most striking difference shared by teleosts studied which distinguishes them from other non-mammalian vertebrates, is that their class I and class II genes are not linked (Bingulac-Popovic et al., 1997; Clark et al., 2001; Grimholt et al., 2002; Hansen et al., 1999; Naruse et al., 2000; Takami et al., 1997). In the case of zebrafish (*Danio rerio*), the class II gene is on a different chromosome (Bingulac-Popovic et al., 1997). As observed in the shark, the expressed classical class I genes identified are found in close association with antigen processing genes. Many genes, including those involved in antigen processing, that are known as being associated with the human class II region, such as TAP2, tapasin, LMP-2 and 7, and RING3, are linked with the class I genes in the Japanese pufferfish (*Fugu rubripes*) (Clark et al., 2001), medaka (*Oryzias latipes*) (Matsuo et al., 2002) and zebrafish (Takami et al., 1997). In Atlantic salmon (*Salmo salar*) (Grimholt, 1997), rainbow trout (*Oncorhynchus mykiss*) (Hansen et al., 1999), and zebrafish (Sultmann et al., 2000), two TAP2 genes have been identified but only one is polymorphic and linked to the MHC class I region. In medaka

and zebrafish, the class I and antigen processing genes are in an uninterrupted cluster. This organization, because of the large evolutionary distance between these fish, could therefore represent the ancestral organization in teleosts (Matsuo et al., 2002).

The number of class I genes in the class I region varies between species. Nine class I loci in Japanese pufferfish have been identified and are linked to the antigen processing genes in the class I region, but their expression has not been characterized (Clark et al., 2001). In the class I region of medaka, two classical and two non-classical class I genes were identified (Matsuo et al., 2002) whereas two or three class I loci are present in zebrafish (Bingulac-Popovic et al., 1997; Takeuchi et al., 1995). Cod (*Gadus morhua*) (Persson et al., 1999) and cichlid fishes (Sato et al., 1997) seem to express a larger number of class I loci. However, the expression of class I in salmonids (Aoyagi et al., 2002; Grimholt et al., 1993; Shum et al., 2001), guppy (*Poecilia reticulata*) (Sato et al., 1995), the channel catfish (*Ictalurus punctatus*) (Antao et al., 1999), and carp (*Cyprinus carpio*) (van Erp et al., 1996) suggests only one classical class I locus is expressed to a significant level despite the presence of more than one locus. Interestingly, the major class I locus in carp seems to be most similar to loci from zebrafish and Atlantic salmon than to alleles at other carp class I loci (van Erp et al., 1996). Some alleles at the expressed class I locus in rainbow trout are also more closely related to alleles from zebrafish and carp (Aoyagi et al., 2002).

The extant fish species, *Latimeria chalumnae,* belongs to the order Coelacanthiformes and is most closely related to the tetrapods (Nelson, 1994). Little research has been done on the MHC of *L. chalumnae* because of the lack of available samples. What is known is that there are a number of MHC class I loci, of which some

are likely classical and expressed (Betz et al., 1994) and are most closely related to *Xenopus* MHC class I.

1.3.3 Amphibians

Amphibians are the most recently derived of the ectothermic vertebrates, and are the oldest of the tetrapods (Kumar and Hedges, 1998). Research on the American axolotl (*Ambysoma mexicanum*), has shown that a number of classical and nonclassical class I genes reside at the MHC class I locus (Sammut et al., 1999). Much more is known about the MHC in different species of African clawed frog, *Xenopus*. The organization of the MHC as found in *Xenopus* gives more evidence to what the MHC was like prior to the evolution of the lineages leading to the reptiles, birds and mammals. The fact that the *Xenopus* MHC has an organization similar to the shark MHC, provides more evidence that the loss of class I and class II linkage in the teleost fish likely occurred in a teleost ancestor (reviewed in Flajnik et al., 1999 and Flajnik and Kasahara, 2001).

The African clawed frog possesses one classical MHC class I gene (Shum et al., 1993) sandwiched between antigen processing genes that are typically known as being located in the mammalian class II region (Nonaka et al., 1997). The class I locus in *Xenopus* contains two ancient allelic lineages at the class I locus (Flajnik et al., 1999a) as can be identified at the dominantly expressed class I locus in sharks and teleosts. The antigen processing genes that are linked to the class I locus, LMP7 (Namikawa et al., 1995), and TAP2 (Ohta et al., 1999) are also composed of two ancient lineages (Nonaka et al., 2000; Ohta et al., 1999).

An interesting characteristic of *Xenopus* MHC is that despite the polyploid genome of many species, a process of diploidization has resulted in only one functional

MHC class I gene expressed. Some members of the polyploid species show expression of a class I locus not linked to the rest of the MHC, while the linked gene is silenced (Flajnik et al., 1999b). Also expressed, but not genetically linked to the *Xenopus* MHC, even in species of *Xenopus* that have not undergone polyploidization, is a large family of non-classical MHC class I genes (Flajnik et al., 1993). Flajnik et al. (1993), hinted that some of the numerous class I genes observed in teleosts, and a region possessing non-classical class I genes in chicken (*Gallus gallus*), could be representatives of this family of genes, but further mapping of the region in the respective genomes is needed to confirm this.

1.3.4 Chicken *(Gallus gallus)*

The MHC of the chicken is the best-characterized avian MHC, and for a long time, was the MHC best characterized in non-mammalian vertebrates. The region containing the classical MHC loci is a compact region on microchromosome 16 centromeric of the nucleolar organizer region (Guillemot et al., 1988; Bloom and Bacon, 1985). The class I (B-F), class II (B-L), and class III regions are in a different order than what is typically observed in mammals. The region containing two classical class I genes is sandwiched between the region containing two class IIβ chain genes (centromeric) and the class III region (telomeric) (Kaufman et al., 1999b) (Figure 4). Altogether, the chicken MHC is about 92 kb in length, approximately 20 times smaller than the MHC class I and class II regions in humans (Kaufman et al., 1999b).

Chicken class I, II, and III gene regions do not possess the entire complement of immune function genes typically associated with these regions in mammals. The only

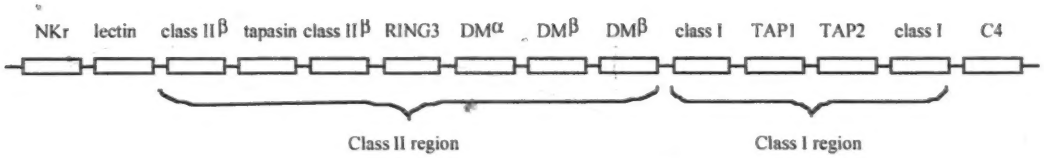


Figure 4: Organization of the MHC in the chicken (*Gallus gallus*) (not to scale). The region spans approximately 92 kb. Adapted from Kaufman et al. (1999b).

class II α gene present in chickens is not associated with the class II region (Kaufman et al., 1995) and the class III region has only been shown to contain one innate immunity gene, identified as complement component C4.

There are few genes in the chicken class I region, with only two classical MHC class I genes flanking TAP1 and TAP2. The TAP genes are located in the class II region in humans and mice, not linked to the classical class I genes as in chicken. In many animals, TAP genes are found in association with other antigen processing genes such as the proteasome associated LMP genes, which have not been identified in the chicken. This finding is consistent with the lack of antigenic peptides that possess hydrophobic or basic carboxyl termini in the ER of some breeds of chicken (Kaufman et al., 1995).

Only the class I gene adjacent to TAP2 is expressed to any significant level on the surface of cells in chickens. This appears to be because the promoter of the class I gene adjacent to TAP1 has either diverged or been deleted (Kaufman et al., 1999a). Recall the association of particular TAP2 and MHC class I alleles in the rat. Kaufman and colleagues (Kaufman et al., 1999a) have proposed that a similar situation has occurred in chicken. Since the TAP and MHC class I genes are so tightly linked, genetic hitchhiking or co-evolution could have allowed the peptide binding preferences of the major MHC class I locus and the TAP genes to be maintained by selection if they were functionally important. Thus, the other (minor) locus may have lost its viability if selection was preferentially maintaining the linkage of the better functioning MHC class I locus and TAP.

The proximity of the B-F/B-L regions means there is also a low level of recombination between the genes of the class I and class II regions (Skjodt et al., 1985).

This can lead to the development of B-F/B-L haplotypes over time. Resistance or susceptibility of certain chicken breeds to Marek's disease virus (MDV) (Briles et al., 1977; Pazderka et al., 1975) and Rous sarcoma virus (RSV) (Schierman et al., 1977) has been correlated with the MHC haplotypes (Kaufman et al., 1999a; Kaufman et al., 1995).

Marek's disease virus is a herpes virus that causes lymphomas. Using chickens bred to be congenic (having identical genes at all loci except for the MHC), different studies related MDV resistance, moderate resistance, or susceptibility with specific MHC haplotypes (Plachy et al., 1992). Recent research shows that resistance is negatively correlated with the level of expression of MHC class I molecules on the surface of cells (Kaufman et al., 1995).

Rous sarcoma virus is a retrovirus that induces rapid tumour growth due the presence of the *v-src* gene, a mutated form of the cellular proto-oncogene *c-src* (Takeya and Hanafusa, 1983). Congenic chickens strains bearing the MHC haplotype B12 are able to survive an RSV infection, whereas those bearing the B4 haplotype do not (Plachy, 1984). Resistant haplotypes express MHC class I molecules that can bind peptides specific to *v-src*, whereas the susceptible haplotypes are not able to bind peptides from *v-src* (Kaufman and Venugopal, 1998).

Other regions in the chicken genome possess MHC related genes that may be linked to disease resistance. The restriction fragment pattern Y (Rfp-Y) region, telomeric to the MHC, contains non-classical MHC genes in the chicken (Afanassieff et al., 2001; Zoorob et al., 1993). Rfp-Y is separated from the classical loci by the nucleolar organizing region, which causes a high level of recombination between the classical and non-classical genes (Miller et al., 1996). The Rfp-Y region has yet to be mapped, but it is

known that it contains at least three MHC class IIβ genes (Zoorob et al., 1993), two MHC class I genes, and a multi-gene family of C-type animal lectins (Bernot et al., 1994; Rogers et al., 2003). One of the class II genes is a pseudogene (Rogers et al., 2003), and only one of the MHC class I genes in the Rfp-Y is expressed, but does not possess some of the classically conserved antigen binding residues (Afanassieff et al., 2001; Rogers et al., 2003) and its function remains unknown. Another MHC class II β locus, not linked to any previously characterized MHC loci, was found by analysis of RFLPs of different haplotypes (Juul-Madsen et al., 2000).

The lack of linkage of the Rfp-Y region with the class I and class II genes of the B-F/B-L region means the cross-breeding of different MHC congenic strains can lead to association of different Rfp-Y regions with different B-F/B-L haplotypes. Research has linked particular Rfp-Y haplotypes to increased resistance to MDV (Sakenell et al., 1996) or RSV (LePage et al., 2000).

In humans and mice, NK receptor genes and the MHC are located on different chromosomes (Yokoyama, 1999). A family of lectin-like genes in the chicken has been identified near the MHC class I genes in the Rfp-Y (Bernot et al., 1994) and lectin-like NK cell inhibitory receptors have been found adjacent to the class II region genes (Kaufman et al., 1999b). Also, a region containing genes that are found in the human NK complex has been identified on chicken chromosome 1 (Bumstead, 1998).

Although the exact function of the lectin-like NK receptor genes in the chicken has not been identified, it is likely that linkage with MHC class I genes has caused them to evolve together. Kaufman (Kaufman et al., 1995) proposes that the linkage of the MHC class I genes with NK cell receptors has allowed certain haplotypes to develop

resistance to some diseases. For instance, chicken strains that show low cell surface expression of MHC class I have increased resistance to Marek's disease virus and those that have a normal expression of MHC class I are susceptible. We know that some NK inhibitory receptors respond to the normal global presence of MHC class I molecules by preventing NK cell killing and that the absence of MHC class I molecules can cause activation of NK cells. This could explain how MHC class I expression has allowed some strains to become resistant to diseases.

1.3.5 Other birds

There are few studies concerning the MHC in other avian species. The quail (*Coturnix japonica*) is the only bird for which a preliminary organization of the MHC region is available for comparison with the chicken (Kulski et al., 2002). The overall arrangement of the quail MHC region is similar to chicken; the antigen processing genes TAP1 and TAP2 are flanked by MHC class I loci and class II genes are nearby with the BG-like and NK-like genes. However, there are seven class I and seven class II β-chain loci in this region in the quail, compared to two of each in the chicken. Other genes that are part of gene families in the quail but are present in single copies in the chicken MHC, are the BG-like genes, and the NK lectin-like and NK inhibitory-like receptors.

So far, only limited studies on the function and expression of quail MHC class I genes have been reported. Four different MHC class I cDNA transcripts were identified (Shiina et al., 1995; Shiina et al., 1999a). Some of the transcripts may not be classical class I genes though (Kaufman, 1999), since genomic sequences identical to two of the transcripts lack some introns (Shiina et al., 1999b). As well, loci corresponding to these

two transcripts are not located in the current genomic map of the quail MHC (Kulski et al., 2002).

The expression of MHC class I and class II genes have been examined in the great reed warbler (*Acrocephalus arundinaceus*) (Westerdahl et al., 1999; Westerdahl et al., 2000). Eight and seven transcripts were identified for class I and II respectively, suggesting the MHC in the great reed warbler is more varied than the chicken with expression of at least four loci for each gene family. However, the genomic location and relative levels of expression of the different loci have not been reported. These important pieces of information are necessary in order to determine the relationship of the loci to that of the chicken.

MHC class II has been analyzed in a few other avian species. The expression of two MHC class II transcripts has been found in pheasants (*Phasanius colchicus*) (Wittzell et al., 1994) and are linked with MHC class I in the MHC proper (Wittzell et al., 1995). The existence of class II genes in an Rfp-Y-like region in the pheasant has also been demonstrated (Wittzell et al., 1995).

An RFLP analysis of MHC class II in the Savannah sparrow (*Passerculus sandwichensis*) reveals a large number of hybridizing bands (Freeman-Gallant et al., 2002) suggesting that there is a large MHC class II gene family. They also found, from the segregation of hybridizing bands among related sparrows, there are two independently assorting clusters of MHC class II. This means there could also be an Rfp-Y-like cluster of MHC genes in the sparrow. Analysis of three amplified products that span a small portion of the Savannah sparrow MHC class II antigen-binding site showed two of the sequences could be alleles at a locus. However, the third sequence only shared

77% identity with them. This third sequence did share 96% identity with a sequence from a bird that belongs to the same family, the red-winged blackbird (*Agelaius phoeniceus*) suggesting that the genes are orthologous (Edwards et al., 1999; Freeman-Gallant et al., 2002). However, more information on the genomic sequence and position of these loci would be necessary to determine their relationship definitively.

Two MHC class II containing cosmids from the red-winged blackbird have been analyzed that span a region almost as big as the entire chicken MHC, (Edwards et al., 2000; Edwards et al., 1998; Gasper et al., 2001). They mapped three class II loci, two of which could be functional, and two additional class II gene fragments (Gasper et al., 2001). It is unknown whether or not these cosmids correspond to the MHC proper, or an Rfp-Y-like region. They show a level of gene density similar to the chicken, but they also possess a variety of retro-elements and evidence of a recent gene duplication, which are not observed in the chicken MHC.

Class II fragments have also been amplified and sequenced from a variety of other birds, such as the great snipe (*Gallinago media*) which is a wading bird (Ekblom et al., 2003), the house finch (*Carpodacus mexicanus*) (Hess et al., 2000), penguin species (*Spheniscidae*) (Tsuda et al., 2001) and a variety of other songbirds (Edwards et al., 1999). Some of these studies confirm what has been seen in many birds; that there are a large number of class II loci, and the organization of the regions surrounding these loci, and the lengths of the non-coding portions of the loci themselves, are very different from those seen in the chicken.

1.4 Rationale for the analysis of MHC in ducks

The White Pekin duck (*Anas platyrhynchos*) is a domesticated breed of mallard. Ducks (order *Anseriformes*) are thought to occupy a position basal to neognathous birds (van Tuinen et al., 2000). They are believed to have diverged earlier from the same lineage (Galloanseriformes) that gave rise to the *Galliformes*, which includes chicken and quail. Given the older evolutionary position of the duck, the MHC of the duck may allow insight into the ancestral avian MHC organization. So far, the chicken MHC is the most extensively characterized within the avian lineage. The MHC of the closely related quail (Shiina et al., 1995; Shiina et al., 1999a) possesses a greater diversity of MHC class I loci (Kulski et al., 2002), similar to more distantly related birds studied (Westerdahl et al., 1999). However, the relative level of expression of MHC class I alleles has not been analyzed in quail or other birds.

Ducks are also of interest because two viruses that are of medical relevance, hepatitis B (HBV) and influenza A, infect them. There are few animals that are naturally infected by hepatitis B. Ducks are congenitally infected by duck hepatitis B virus (DHBV) (Mason et al., 1980; Urban et al., 1985) and are well developed as an animal model for HBV research. Until duck hepatocytes were used (Tuttleman et al., 1986), *in vitro* studies of hepatitis B were not possible and live animal research was costly.

Ducks are considered to be the primary, primordial reservoir of Influenza A virus strains (Webster et al., 1992). Avian influenza strains infect the intestine of the duck, and there are no clinical signs of disease in the animals except that virus is shed in high titres in their feces (Kida et al., 1980; Webster et al., 1978). This allows easy transmission of viral strains to other individuals since the virus is still viable from the feces after a week

at 20°C, or up to a month at 4°C (Webster et al., 1978). Influenza A strains causing deadly outbreaks in humans, such as the 1918 pandemic (Gorman et al., 1991) or the Hong Kong epidemic in 1968 (Laver and Webster, 1973), originated in ducks. There is a high degree of genetic similarity (or a lack of mutation) in the avian viruses circulating now compared to the strains that caused human illness, indicating that there is little selection on the virus by the duck host, and that the original strains came from ducks (Webster et al., 1992). We are interested in knowing if there are characteristics of the duck MHC that could influence the immune response to these viral challenges.

1.5 Hypothesis and Objectives

Our hypothesis was that ducks would have characteristics of the MHC similar to chicken. Specifically, we thought that: 1. Ducks would possess a small number of MHC class I loci; 2. Ducks would show dominant expression of MHC class I from a single locus; 3. The locus encoding the dominant MHC class I would be adjacent to TAP2. At the start of the project, the sequence encoding the MHC class I of ducks was not known.

The first objective was to determine the number of MHC class I sequences. To do this, we made a cDNA library from duck spleen and screened the unamplified library for MHC class I. Because the library was not amplified, the number of clones containing a sequence would reflect the relative level of expression of that sequence. We also amplified MHC class I from genomic DNA to identify other loci not represented in the library and to observe MHC class I diversity in the genome.

The polymorphism of TAP2 in chickens led us to our second objective of screening the cDNA library for TAP2. Genomic PCR amplification and RT-PCR were also employed to look for TAP2 allelic polymorphism. In chickens, the polymorphism is a result of TAP2 linkage with the dominantly expressed MHC class I locus. Our third objective was to analyze the genomic organization of MHC class I and TAP2 by amplifying the region between these two genes.

2 METHODS

2.1 Laboratory ducks

White Pekin ducks were kindly provided by Dr. D. L. Tyrrell (University of Alberta) from the breeding stock of his duck colony. Animals were euthanized with Euthanol and organs and blood were isolated. Organs were frozen immediately in liquid nitrogen and stored at –80°C while blood samples were treated with heparin and kept at 4°C prior to isolation of DNA.

2.2 Isolation of high molecular weight DNA

Heparinized blood samples were washed in phosphate buffered saline (PBS) and suspended in 1X SSC [150mM NaCl and 15mM Sodium Citrate, pH 7]. Cells were lysed by addition of an equal volume of TNE buffer [100mM Tris (pH 7.5), 100mM NaCl, 10mM disodium ethylenediaminetetraacetate (EDTA) and 1% w/v Sarkosyl] and the proteins digested overnight at 55°C with 100 µg/ml proteinase K. The DNA was gently extracted twice with an equal volume of phenol, followed by a 1:1 mixture of phenol and chloroform:isoamyl alcohol (24:1) and finally, chloroform:isoamyl alcohol (24:1). The DNA was dialyzed for 24 h against 4L of TE [10mM Tris (pH 8), 1mM EDTA (pH 8)] with 100mM NaCl and for another 24h against 4L of TE.

2.3 Generation of the MHC class I probe (D26E4p)

Degenerate primers CHE4F1 (Table 1) and CHE4R1 were designed based on alignment of the exon 4 sequences of chicken (Hunt and Fulton, 1998; Kroemer et al., 1990), quail (Shiina et al., 1995) and crane (Jarvi et al., 1999). These primers were used to amplify a 228 bp fragment from the genomic DNA of five ducks (ducks #26, 95, 105, 129, and 132), using an annealing temperature of 54 °C with standard PCR conditions.

*Table 1: Oligonucleotides used for the amplification and sequencing of A. MHC class I and B. TAP2 products.

A. MHC class I primers

Primer Name	Primer sequence	Description
CH E4 F1	CCCGAGGTSCGAGTGTSG	Chicken, forward MHC class I Exon 4
CH E4 R1	CACGCGGCACYGGTACTTGTC	Chicken, reverse MHC class I Exon 4
5' UT F2	AGGCAGCTCGGGGCACAG	Duck 26, MHC I before signal peptide
D26 SP1	CTCGTCCTGGGGCTGCTGCTG	signal peptide primer, MHC class I, few mismatches with clones U*03 and U805
D26 SP2	GGCTGCTGCTGGGGGTCCTG	signal peptide primer, MHC class I, that matches four cDNA transcripts
D26E2F1	GCCCCACTCCCTGCGCTA	Duck, fwd, MHC I, exon 2 start
D26E2F2	CGGGAGCGCTACAACCAGAG	Duck, fwd, MHC I, exon 2 end
D26E2R1	CTCTGGTTGTAGCGCTCCCG	Duck, rev, MHC I, exon 2 end
D26PE3F1	ATGTATGGCTGTGACCTCCTC	MHC class I, exon 3 fwd (mismatch with U*02 and U*03)
E3F2	CCTGGAGAACACCTGCATTGGG	Duck, cDNA MHC I, mid exon 3, all clones, FWD
D26 E3F3	ATGTATGGCTGTGACCTCCT	Duck, exon 3 fwd (better match to all clones)
D26PE3R1	GAGGAGGTCACAGCCATACAT	MHC class I, exon 3 rev (mismatch with U*02 and U*03)
D26E4F1	CCGCGACCCATCTCCATC	MHC class I, exon 4 forward, match clones
D26E4R1	GGAGATGGGTCGCGGGTAG	MHC class I, exon 4 forward, match clones
D26E4R2	GTGAGCACGGCAGGACAAGGT	Duck, rev, MHC I, exon 4 mid
TMF	GCTGCCCTGGCTGGATTTGC	Duck 26, end of TM MHC I, FWD
CY1 REV	TTGTAGCCCTTCCCCTTCTTC	End of CY1, MHC class I, reverse, matches four cDNA transcripts
CY3 Rev	GGTTAGACACTGGGGTTGCTCC	Duck 26, rev of UTF MHC I (last CY and beg. 3' UTR
D26 3'UTR Rev	GCAGATAGGAGATGTGAGAGGTTG	3'UTR, MHC class I, matches expressed loci (NOT U*05)
D26 5-1 UTF	GGAGCAACCCCAGTGTCTAACC	Duck, for U*04 3'UTR MHC I, fwd, matches all clones
D26 5-1UTR	CCCCAATCAGACTGCTCAAATC	Duck, for U*04 3'UTR I, rev, matches all
c6-13 3'UTF	TCACGAGAAAGAGAATCACAGAGG	Duck, cDNA MHC I clone U*04 3' UT forward
c6-13 3'UTR	CACCAGGCCAGGATAGGAGCA	Duck, cDNA MHC I clone U*04 3' UT reverse (double primes)
c3-24 3'UTR	GACCAACCATTACCACGACAAC	Duck, cDNA MHC I clone U*05 3'UT rev
c3-24 3'UTF	TGCTGCTTCAATGTCCTGGTTTC	Duck, cDNA MHC I clone U*05 3' UT fwd
c3-24 3'UTF2	AGGTGACCCAACTAGCCAAAG	Duck 26, cDNA MHC I clone U*05 3'UT, further forward (closer to poly AAA)
c2-20 3'UTF	ATCCGCAGCGAGCCCTTCCAC	D26, clone 2-20 (matches U*05) amplify after first string of A's (3' UTR)
c2-20 3'UTR	AGAGCGATTGGGAAGCATGAG	D26, clone 2-20, reverse from 3'most end (double primes)

* Table I continued

Table 1: B. TAP2 Primers

Primer Name	Primer Sequence	Primer Description
CHT2E8R1	CGATGCGCTGCTTCTGCCC	Chicken, reverse Exon 8 #1
CHT2WAF	AATGGCAGCGGGAAGAGCAC	Chicken, forward Walker A
CHT2WBR	TGGTGGCTTCGTCGAGGATA	Chicken, reverse Walker B
DT2E7F	CCATCCGGGACAACATTGCCTAT	Duck, forward Exon 7
DT2E6R	TCGAAGATTTTGTGGGCAGCTG	Duck, reverse Exon 6
CHT2E15UF	GGATCCCCACCCCACAGCCATG	Chicken, forward Exon 1 5'UTR
CHT2E1F1	CACGCTGCTCCTGGCCGACCT	Chicken, forward Exon 1 #1
CHT2E1F2	CCGACCTGGCCCTCATGTTGG	Chicken, forward Exon 1 #2
CHT2E1F3	CCCTGGCCCACTTCTTCCC	Chicken, forward Exon 1 #3
CHT2E1F4	CCGCCATCTTCCTGACCCTACG	Chicken, forward Exon 1 #4
CHT2E1F5	CCTGGAGCTGCCCCGGTGC	Chicken, forward Exon 1 #5
DT2WAF	TGAACGGCAGTGGGAAGAGCA	Duck, forward Walker A
DT2E8R1	CAACACGCTGCTTCTGCCCA	Duck, reverse Exon 8 #1
DT2E1F1	TGAGACCTCCGTGCCCTACTG	Duck, forward exon1 #1
DT2E1F2	CACATGGCTGGAGGCTGGG	Duck, forward Exon1 #2
DT2E4R	CCCGCCGGATGAGGAGGA	Duck, reverse Exon 4
DT2E2F	CTTCCAGCAGACCACAGCAGG	Duck, forward Exon 2
DT2I3F	TGGTGGACATGGGACAGGAGAA	Duck, forward Intron 3
D9T2I3R	TTCCTGCTCTACTCTTGCCTGTT	Duck, reverse Intron 3
DT2E1R	CCAGCCTCCAGCCATGTG	Duck, reverse Exon 1
DT2E3R	CAAGCAGCGCCAGCAGGGTCAG	Duck, reverse Exon 3
DT2E4F	GGTGCAGGAGGCCATCT	Duck, forward Exon 4
DT2E1R1	GATGGCAGTGGCAAAGGCAGTG	Duck, reverse Exon 1 #1
DT2E93UR	TCAGCTGCTCCATGGGTTGTTAG	Duck, reverse Exon 9 and 3' UTR

The product was gel purified using the QiaQuick Gel Extraction Kit (Qiagen, Mississauga, ON) and cloned with the TOPO TA cloning kit into the PCR 2.1 TOPO vector (Invitrogen, Carlsbad, CA) according to the manufacturer's directions. Clones were sequenced and aligned with avian MHC class I sequences. The PCR product from each duck consisted of several MHC class I sequences.

2.4 Generation of the TAP2 probes

Kyle Thulien designed the probe for TAP2 and screened the amplified cDNA library with it. The probe was amplified from duck genomic DNA using primers based on the Walker A motif located in exon 6 and the Walker B motif in exon 8 of the published TAP2 sequences of chicken and quail. The primers CHT2WAF and CHT2WBR amplified a 475 bp fragment using standard PCR conditions with an annealing temperature of 53°C. The product was gel purified, sequenced, and determined to be homologous to chicken and quail TAP2 sequences.

Kyle Thulien performed the Southern hybridizations with the TAP2 probe described below. For Southern hybridization, a probe (tap2WA8p) was generated using the cDNA clone 6.2.2 as a template with the primers DT2WAF and DT2E8R1. A 335 bp product (from exon 6 to exon 8) was amplified with an annealing temperature of 65°C using the standard PCR conditions. For hybridization, probes were radiolabelled with ^{32}P α-dCTP by random priming (Feinberg and Vogelstein, 1983) (PrimeIt Random Primer Labeling Kit, Stratagene, La Jolla, CA).

2.5 cDNA library construction and screening

Messenger RNA (mRNA) was isolated from the spleen of duck #26 (D26) using the FastTrack™ 2.0 kit (Invitrogen, Carlsbad, CA) according to the manufacturer's

protocol. Five micrograms of mRNA was used to construct a cDNA library (ZAP Express® cDNA Synthesis Kit and ZAP Express® cDNA Gigapack® III Gold Cloning Kit, Stratagene, La Jolla, CA). One third of the library (100 000 plaques) was plated on XL1-Blue-MRF' cells, as an unamplified library, and duplicate Nytran Supercharge nylon transfer membranes (Schleicher and Schuell, Keene N.H.) were lifted and hybridized with the MHC class I probe D26E4p. After hybridization, filters were washed three times at 52°C for 20 min each in 1X SSPE [150mM NaCl, 10mM NaH_2PO_4-H_2O, 0.1 mM EDTA] and 0.1% w/v SDS. Two hundred hybridizing plaques were isolated. Fifty-six of these were taken through two additional rounds of hybridization to isolate positive clones. Thirty-one clones were converted to plasmids and sequenced.

The other two-thirds of the library was plated on host cells and eluted in SM buffer [0.1M NaCl, 0.8M Magnesium sulfate, 0.5 M Tris (pH 7.5), 0.1% w/v gelatin] and stored at 4°C with chloroform (0.3% v/v). An aliquot of this once amplified library was screened for TAP2 with the probe tap2WABp.

2.6 PCR based sequencing of positive plaques

For 32 of the positive plaques, we amplified exon 2 under standard PCR conditions using the primers D26E2F1 and D26E2R1 at an annealing temperature of 55°C. The 257 bp product was extracted from the gel by freezing the gel fragment containing the band of interest under liquid nitrogen, centrifuging the frozen fragment for 10 min, and extracting the liquid phase containing the DNA from the powdered agarose. The DNA was then purified and the nucleotide sequence determined.

2.7 RT-PCR amplification of MHC class I and TAP2 alleles

We synthesized first strand cDNA from 500 ng of D26 spleen mRNA by priming with LNK-Tail (5'-TCTGAATTCTCGAGTCGACATCTTTTTTTTTTTTTTTT-3') and extending with the Thermoscript RNase H- Reverse Transcriptase kit according to manufacturer's instructions (Invitrogen, Carlsbad, CA). The MHC class I forward primers D26SP2 or D26E2F1 and the reverse primer (AdapTTT: 5'-TCTGAATTCTCGAGTCGACATC-3') were used to amplify MHC class I using standard PCR conditions and an annealing temperature of 60°C. The primers DT2E4F and DT2E8R1 were used to amplify in standard conditions exon 4 to exon 8 of TAP2 using a 1 in 100 mix of Pfu in Taq polymerase (Barnes, 1994) with an annealing temperature of 63°C. The products were gel purified, cloned into the Topo-TA vector and individual clones sequenced.

2.8 Amplification of genomic clones of MHC class I

Genomic fragments of MHC class I from D26 were first amplified from the 5' untranslated region (primer 5'UTF2) to exon 2 (primer D26E2R1), exon 2 (D26E2F1) to exon 3 (D26E3R1), and exon 3 (D26E3F1 or D26E3F2) to exon 4 (D26E4R2). Standard PCR conditions were used with an annealing temperature of 55°C. Amplified products were band purified by gel electroelution (Sambrook et al., 1989), cloned into the Topo-TA vector and sequenced.

Longer genomic MHC class I sequences were amplified with a primer in exon 1 (D26SP2) and a primer in exon 6 (D26CY1) using the conditions for amplification of longer PCR products, at an annealing temperature of 63°C. Amplified products were

purified (QIAquick PCR purification kit Qiagen, Mississauga, ON), cloned into the Topo-TA vector and sequenced.

2.9 Amplification of genomic clones of TAP2

Kyle Thulien originally amplified TAP2 from genomic DNA. This was redone, to confirm the sequence, by amplifying a variety of overlapping genomic TAP2 PCR products. Exon 1 to exon 8 was amplified (primers DT2E1F2 and DT2E8R1) using the amplification conditions for long PCR products. Other amplifications were performed using standard PCR conditions. Exon 1 to exon 4 (primers DT2E1F2 and DT2E4R) was amplified using a 1 in 100 mix of Pfu in Taq polymerase (Barnes, 1994), and an annealing temperature of 63°C. Exon 4 to exon 8 was amplified using an annealing temperature of 60°C (primers DT2E4F and DT2E8R1). Intron 3 (DT2I3F) to exon 9 (DT2E93UR) was amplified with an annealing temperature of 55°C. Finally, exon 1 (CHT2E1F1) to exon 6 (DT2E6R) was amplified with an annealing temperature of 60°C. Amplified products were purified, cloned into the Topo-TA vector and sequenced and a contig of each TAP2 allele was generated.

2.10 Amplification of the intergenic region of MHC class I and TAP2

We amplified the genomic region between the 3' end of an MHC class I locus and the 3' end of the TAP2 locus from D26 and D95. The primers D26TMF and DT2E7F were used under standard PCR conditions with an annealing temperature of 55°C. The product was cloned into the Topo-TA vector, sequenced by primer walking, and a contig of overlapping clones was generated.

2.11 Southern hybridization

High molecular weight DNA (10µg) from erythrocytes of White Pekin ducks and chickens was digested to completion with 40 U of restriction enzymes for 5 hours and then with another 40 U overnight, according to the manufacturer's protocol (New England Biolabs, Mississauga, ON). The digested DNA was purified by phenol: chloroform extraction, and ethanol precipitated. DNA was separated on a 0.8% agarose gels in 0.5X TBE buffer for 16hrs at 20 V, blotted onto Nytran Supercharge nylon transfer membrane and immobilized by UV cross-linking (UV Stratalinker© 2400, Stratagene, La Jolla, CA). The blots were pre-hybridized [50% formamide; 5X Denhardts; 4X SSPE; 5% Dextran Sulfate; 1% SDS] at 42°C for 6 hours and hybridized overnight with 100µg/mL salmon sperm DNA and duck specific probes. The blots were washed at high stringency [1X 100 ml: 1X SSPE and 0.1% SDS; 3X 300 ml: 0.1X SSPE; 0.1% SDS] and exposed to Kodak X-OMAT AR scientific imaging film (Rochester, New York) at -80°C.

2.12 Northern hybridization

Total RNA was separately purified with Trizol (Gibco BRL, Burlington, ON) according to manufacturer's instructions, from the spleen, heart, kidney, testis and liver of duck #132. Cynthia Radford performed the Northern blot hybridization. The total RNA from different organs and the mRNA isolated as previously described were run on a 1.2% agarose, 0.6% formaldehyde gel for 30 min at 60 V and then 2.5 hrs at 100 V. The gel was blotted onto Nytran Supercharge nylon transfer membrane and UV cross-linked (UV Stratalinker© 2400, Stratagene, La Jolla, CA). The blot was pre-hybridized at 42°C for 2 hours and hybridized overnight with the MHC class I probe D26E4p and 100µg/mL

salmon sperm DNA. The blot was washed at high stringency and exposed to Kodak X-OMAT AR scientific imaging film at -80°C overnight.

2.13 Polymerase chain reaction conditions

Amplifications done in a GeneAmp PCR system 9700 (PE Applied Biosystems Inc., Foster City, CA) were based on the following standard conditions with annealing temperatures specific for each primer pair used: 94°C for 5 min; 30 cycles of 95°C for 30 sec, (53°C to 65°C) for 30 sec, and 72°C for 1 min 30 sec; 72°C for 5 min. We used 100 ng duck genomic DNA as template and 20 pmol of each primer or 100 pmol if the primer was degenerate. Reactions were performed using 0.2 mM dNTP (Gibco BRL, Burlington, ON) 1X PCR buffer, 1X Solution Q, 1 unit Taq polymerase (Qiagen, Mississauga, ON). Longer PCR products were amplified under the following conditions: 94°C for 1 min; 10 cycles of 94°C for 10 s, 63°C for 1 min, 68°C for 2 min; 25 cycles of 94°C for 10 s, 63°C for 1 min, 68°C for 2 min, with the extension time increasing by 2 s with every cycle; 68°C for 5 min).

2.14 DNA sequencing and analysis

Nucleotide sequences were determined on both strands with the cycle sequencing method using the DYEnamic ET terminator cycle sequencing kit (Amersham Pharmacia Biotech) on an automated ABI 377 DNA sequencing machine (PE Applied Biosystems Inc., Foster City, CA).

All chromatogram editing, nucleotide alignment, primer design, and percent nucleotide identity analysis was performed using Genetool (Biotools, Edmonton, Alberta) and by visual inspection. Protein alignment and analysis was done using Peptool (Biotools, Edmonton, Alberta). Tmpred (Hofmann and Stoffel, 1993) and Tmap (Persson

and Argos, 1994; Persson and Argos, 1996) were used to help predict the topology of Tap2. The program Lalign (http://www2.igh.cnrs.fr/home.eng.html) was used to create the edited alignments presented.

3 RESULTS

3.1 Amplification of MHC class I exon 4 sequences from different ducks

Our initial strategy to identify MHC class I sequences from ducks was amplification from genomic DNA using degenerate primers (CHE4F1 and CHE4R1) which were based on highly conserved regions of other avian MHC class I sequences. These primers amplified a fragment of 189 bp (primers not included). We subcloned these fragments in the TA overhang PCR 2.1 TOPO vector (Invitrogen) and sequenced 10-15 clones from each duck. The sequences corresponded to exon 4 from several putative class I genes. During our initial work a clone encoding MHC class I of ducks was submitted to Genbank (Accession number AF393511). We have named the allele *Anpl*U*01. We translated our amplified exon 4 fragments and aligned them with the exon 4 amino acid sequence from U*01 (Figure 5). Each unique sequence was given an allele designation. The letter g denotes that these are sequences derived from genomic DNA. The letter N denotes the sequences encoding a stop codon.

Each duck carried different MHC class I alleles (Table 2). The greatest number of unique exon 4 sequences was amplified from D26, which yielded 5 sequences: g*Anpl*U*02, gU*04, gU*06, gU*07, and gU*10N. Some alleles were shared between related ducks in the duck colony including sequence gU*02 which was found in D26, D95 and D132. Allele gU*04 was sequenced from both D26 and D129. The sequence gU*10N includes an in frame stop codon, and is carried by all individuals analyzed. Alleles gU*08 from D95, and gU*11N from D132 were not identified in the other ducks.

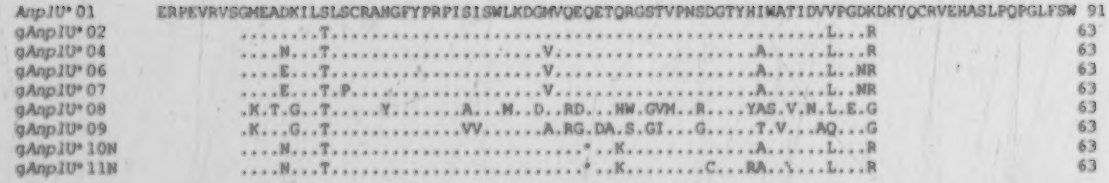


Figure 5: Alignment of the MHC class I exon 4 amino acid sequences amplified from different ducks. The sequences have been aligned to exon 4 of the sequence *Anpl*U*01 (GenBank Accession number AF393511). Identity with *Anpl*U*01 is shown by a *dot*. The genomic sequences have been renamed to correspond with cDNA sequence names. The letter g is included before the name to indicate that it is a sequence derived from genomic DNA. Sequences designated by an N contain a stop codon (*).

Table 2: MHC class I alleles amplified from DNA of five individuals from the duck colony. Alleles are identified on the basis of unique exon 4 sequences. Numbers in brackets refer to the number of clones sequenced from each duck.

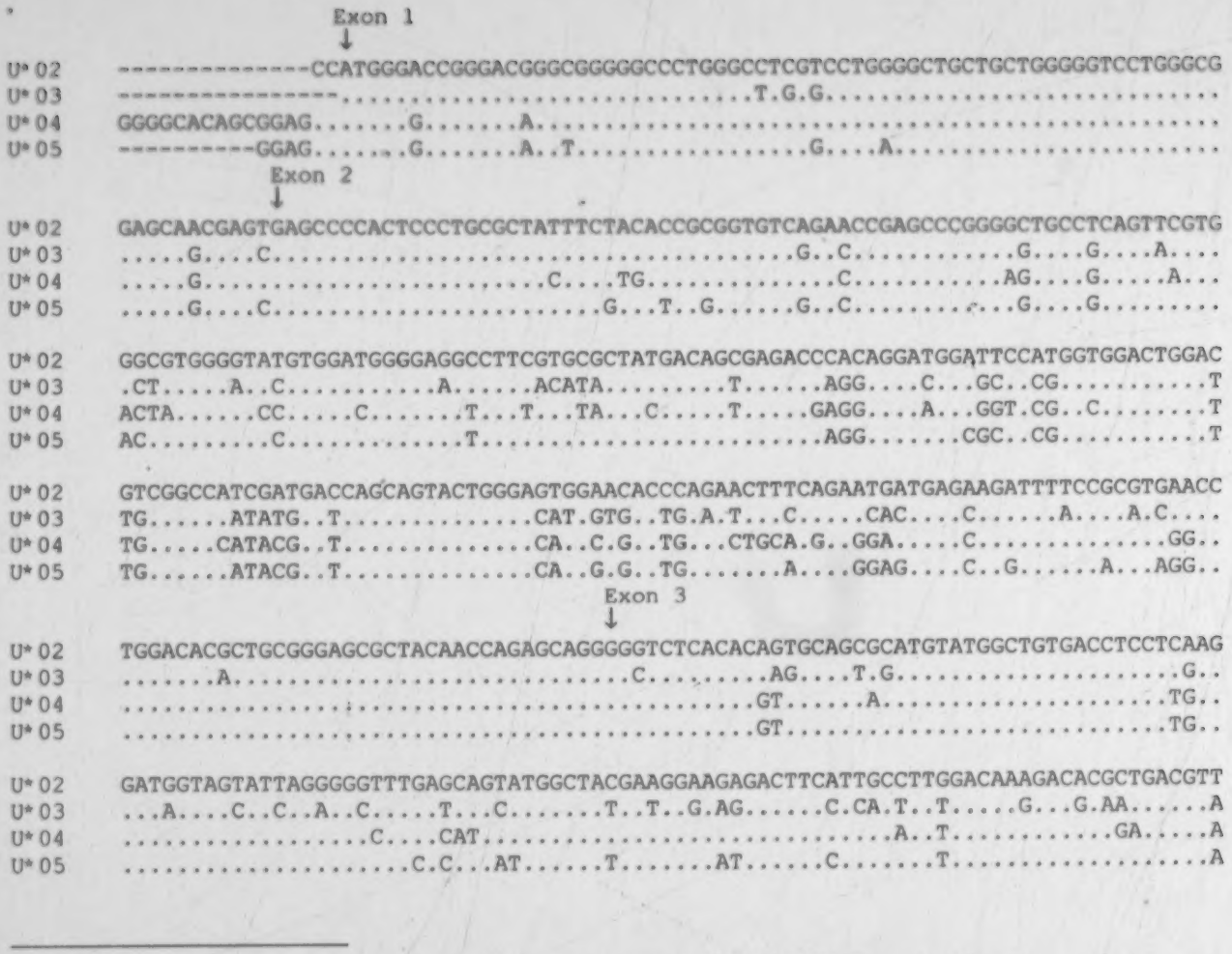
D26	D95	D105	D129	D132
gU*02 (5)	gU*02 (2)			gU*02 (1)
gU*04 (2)			gU*04 (2)	
gU*06 (1)				
gU*07 (1)				
	gU*08 (1)			
	gU*09 (1)			
gU*10N (5)	gU*10N (6)	gU*10N (11)	gU*10N (6)	gU*10N (7)
				gU*11N (4)

3.2 Isolation of MHC class I cDNA clones

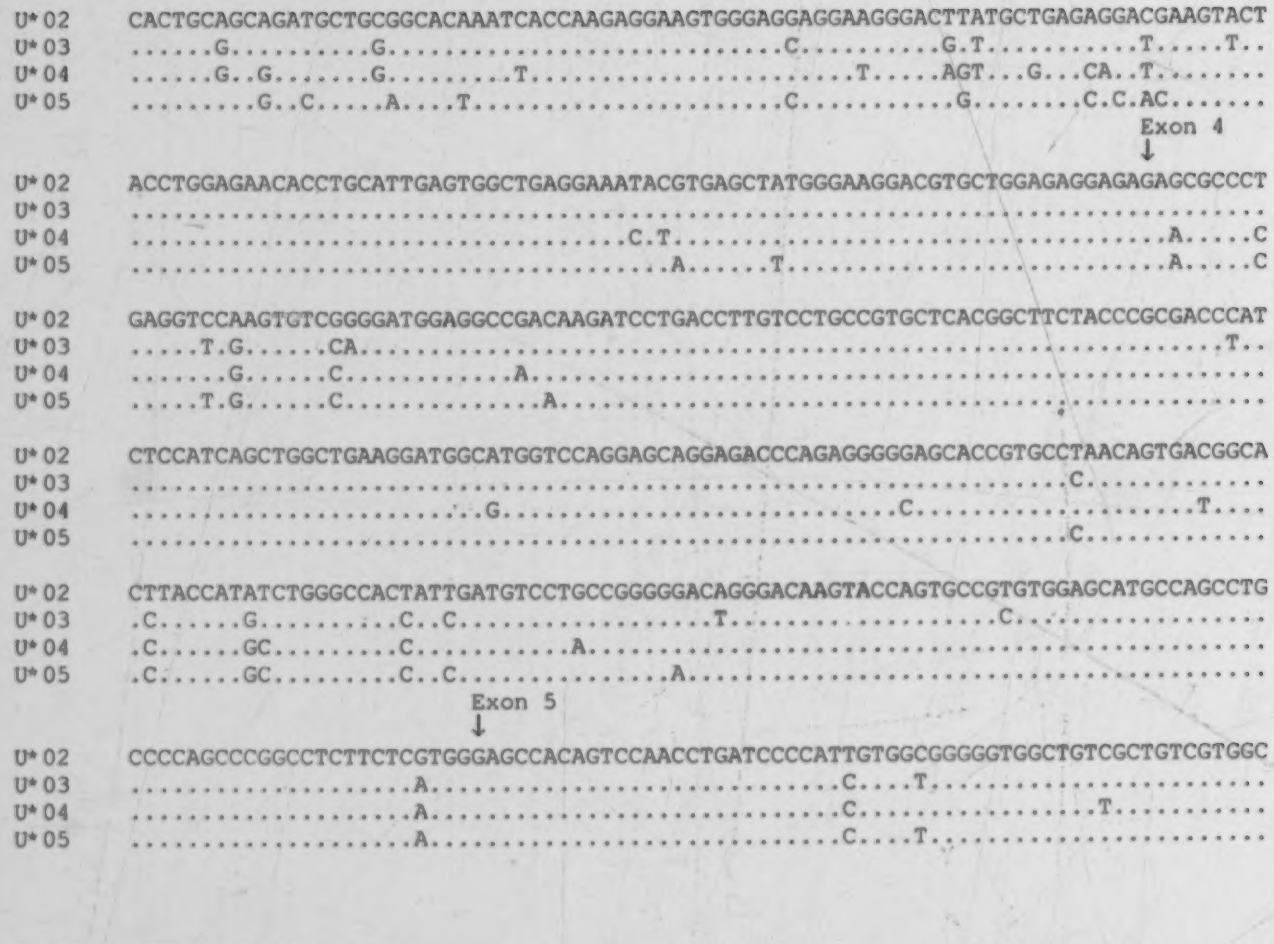
To provide an unbiased representation of the MHC class I sequences expressed by an individual duck, we screened an unamplified duck spleen cDNA library. Since the greatest number of unique exon 4 sequences was amplified from D26, we constructed the cDNA library from the spleen of D26. The PCR product from D26 genomic DNA that contained several MHC exon 4 fragments was used as a probe (D26E4p). We knew D26 was a breeding male from the duck colony. Farmed ducks are deliberately outbred, so D26 was likely a heterozygote. Two hundred positively hybridizing clones were identified in $1x10^5$ clones screened. Thirty-one clones were carried through a third round of hybridizations and plasmids were excised and sequenced in one direction. This sequencing identified only four unique MHC class I sequences, denoted U*02 (22 clones), U*03 (4), U*04 (3), and U*05 (2).

To characterize sequences present in the unpurified phage plaques, we amplified 257 bp of exon 2 directly from the phage supernatant of 13 clones. Analysis of the exon 2 sequences indicated that nine were identical to U*02, three to U*03, and one to U*04. In total, there were 31 copies of U*02, seven of U*03, four of U*04, and two of U*05 identified in the cDNA library. The relative number of clones of each sequence in the unamplified library indicates directly the level of expression of these genes.

We completely sequenced one clone for each of the four unique sequences. The open reading frame (ORF) of the full length clones for U*02, U*03, and U*04 are of identical length (1077 bp) (Figure 6). The sequence for U*05 lacked the third cytoplasmic domain making the ORF for U*05 1030 bp long. Two alternately spliced clones of U*02 lacked the second cytoplasmic domain. The percent nucleotide identity



* Figure 6: Nucleotide alignment of the four duck MHC class I sequences from the cDNA library (continued).



* Figure 6: Nucleotide alignment of the four duck MHC class I sequences from the cDNA library (continued).

* Figure 6: Nucleotide alignment of the four duck MHC class I sequences from the cDNA library (continued).

```
U*05 [ 3' UTR]
        CTGGTGCCCTGTGGAGCGTGTTCCAGTGCCTGACCACCCCGTCAGTGAAGAACCTTTTTGTGATATCAATCCTGAACCAT
        CCCGGTTGCAGCTTCAAGCCATCTCCTCAGGTCCTAATCGCTGGTTGCTGTGAGAATGAAAGTTGTCTCTTTGCAGTAGA
        AACTTCCACTAACCTTGCATATTCAAAAGTGATTGTGCATGCTTAACAGTGGGGGGTATGGTTGTCGTGGTAATGGTTGG
        TCAAGTAAAACTGTACCTACCTGGTATAGTAATAAGTGTCGAGTAGTTTGTGAACCTGTAGCACTCAGCCAATGAGGAAA
        CAGGGGAGGAATCAAGTGTCAGCTAATAGGGAATATAAGGTTATGATGTACTTGCAGCACGTCCGCCATTGCAATCACAA
        AAATGCTGCTTCAATGTCCTGGTTTCAGTTAGAACAGAATTAATTTTCTTCCTAGTAGCTGGTGGAATGCTGTGTTTTGG
        CTTAGGATGAGAAGAGTGCTGATAACACCCCAATGCTTTAATTGTTGCAGAGCAGTGCTTACACCAAGCCAAGGACATCT
        CAGCCTTTTGCTCTGTCCTGCCAACGGGCAGGCTGGGGGTGCAGTAAGAGCTGGGAGGGGACAGACCCAGGACAGGTGAC
        CCAAACTAGCCAAAGGGGTATTCCATACCATCTGACGTCATGCTAAACAATATATAGGGGTGGCTAGCTGGGGGGAGGGG
        GCCGGACTGCTCGGGGTTAGGCTGGGCATCGGTCAGCGAGTGGTGAGCAATTGCATTGTGCATCACTTGTTTGTACATAT
        GATTATTACTTTCCTATTATCACCATTGTATTATTATTATTATTGTTATTATTATTTTTGTTATTATTATTTTCCTG
        TCTTATTAAACTGTCTTTATCTCAAAAA
```

Figure 6: Nucleotide sequences of the four duck MHC class I sequences from the cDNA library are aligned with sequence U*02. Alleles U*02, U*03, and U*04 are aligned up to the first polyadenylation site. The longer 3' untranslated region of U*04 using the second polyadenylation site is below the alignment. U*05 is aligned with the other alleles up to the stop codon, and the 3' untranslated region is given below the alignment. Dots (.) indicate identity with U*02, and spaces, indicated by dashes (-), are included to maximize the alignment. Exon boundaries of the cDNA sequences are indicated.

of the four complete sequences ranges from 88.5% to 91.1% identity (Table 3). Comparing exons of each sequence separately, the percent identity between exon 3 is 80.3% to 90.6%. The percent identity of exon 4, the conserved immunoglobulin domain, is 94.5% to 96.7%. Clones U*02 and U*04 matched genomic sequences amplified from D26 (gU*02 and gU*04, see Figure 5).

We also compared the 3'untranslated region between the clones. The 3'-untranslated region is 151 bp for U*02, 153 bp for U*03 and U*04. A second polyadenylation signal sequence used in U*04 is 814 bp from the stop codon. U*05 had a 3'untranslated region 933 bp long, which ended in a string of 22 As with no polyadenylation signal sequence, indicating internal priming likely. The percent nucleotide identity between the stop codon and first polyadenylation site was calculated. U*02 and U*03 are 82.8% identical, U*03 and U*04 are 82.4% identical, and U*02 and U*04 are 78.2% identical. None of the sequences share significantly greater levels of identity in the coding region or the 3'untranslated region.

Four clones from the library, identical to U*02 in the coding region, carried nucleotide sequence 5' of the start codon. Clone #3-23 had the longest 5' untranslated sequence identified: ggataccttcgagctccggaggcagctcggggcacagcggagcc.

3.2.1 Reverse-transcriptase PCR amplification of MHC class I

To confirm the biased representation of alleles isolated from the cDNA library we amplified MHC class I sequences by reverse-transcription PCR (RT-PCR). First strand cDNA made from D26 spleen mRNA was used as a template for PCR amplification of expressed MHC class I sequences. A forward primer in the exon coding for the signal peptide (D26SP2) or from exon 1 (D26E2F1) and an oligo dT primer (AdapTTT) were

Table 3: Pairwise comparisons of the coding region among five duck MHC class I cDNA sequences. Values above the diagonal are percent nucleotide identity whereas the values below the diagonal are percent amino acid identity between each pair.
*Comparisons involving the cytoplasmic domain are omitted for allele U*05 because it lacks the last two exons of the cytoplasmic domain.

		U*01	U*02	U*03	U*04	U*05
U*01	Entire	———	89.4	88.0	89.1	89.8
	exon 2	———	81.1	83.0	81.8	86.7
	exon 3	———	90.9	83.0	89.4	88.7
	exon 4	———	94.1	94.5	93.8	94.4
	TM	———	93.1	92.2	96.1	93.1
	CY	———	89.3	90.5	85.7	———
U*02	Entire	85.6	———	90.2	89.7	91.1
	exon 2	70.5	———	81.8	80.3	84.9
	exon 3	85.9	———	88.0	90.6	90.6
	exon 4	95.6	———	95.2	94.9	95.2
	TM	97.1	———	97.1	97.1	98.0
	CY	85.7	———	98.8	89.3	———
U*03	Entire	84.8	85.8	———	88.5	91.1
	exon 2	78.4	72.7	———	84.1	87.1
	exon 3	75.0	79.4	———	85.5	84.8
	exon 4	94.5	95.6	———	94.5	96.7
	TM	91.2	94.1	———	96.1	99.0
	CY	85.7	100.0	———	88.1	———
U*04	Entire	85.0	84.4	84.9	———	89.8
	exon 2	73.9	65.9	75.0	———	87.1
	exon 3	83.7	83.7	80.4	———	88.4
	exon 4	93.4	95.6	94.5	———	96.3
	TM	97.1	100.0	94.1	———	97.1
	CY	78.6	82.1	82.1	———	
U*05	Entire*	83.9	85.0	85.6	85.9	———
	exon 2	79.6	70.5	75.0	78.4	———
	exon 3	78.3	84.8	81.5	80.4	———
	exon 4	93.4	95.6	94.5	96.7	———
	TM	94.1	97.1	97.1	97.1	———
	CY*	———	———	———	———	———

used to amplify products of 1211 bp or 1183 bp in length respectively. The products were cloned and sequenced, and 17 sequences were identical to U*02 while 18 were identical to U*03.

3.2.2 Northern hybridization

To determine the level of expression of the MHC class I alleles in various tissues we performed Northern blot analysis (Figure 7). We used the probe D26E4p, which is composed of many MHC class I exon 4 sequences. There was a high level of expression in the spleen and the liver. Heart, kidney and testes showed a significantly lower level of expression. Only one hybridizing band of approximately 1.2 kb was visible. In the mRNA from D26 spleen, the faintly visible larger band of approximately 4 kb corresponded to unspliced pre-mRNA. These results suggested there was significant transcription of class I products corresponding to 1.2 kb in length. Alleles U*02 and U*03 produced transcripts of similar lengths (1159 bp or 1161 bp respectively) whereas U*04 was identified as transcripts of 1161 bp and 1822 bp in length. If allele U*05 had a polyadenylation site beyond the 933 bp 3'untranslated region we sequenced, it would produce a significantly larger transcript, which was not observed. There was no hybridization on the northern blot corresponding to the larger products from U*04 and U*05, suggesting they are transcribed at a very low level.

3.2.3 MHC class I amino acid sequences

To compare the four duck MHC class I sequences we determined the putative amino acid sequence. The deduced amino acid sequences were aligned to compare the structural features of the four cDNA sequences identified (Figure 8). The alignment included the chicken MHC class I allele BF19 and the quail *Coja*-B allele.

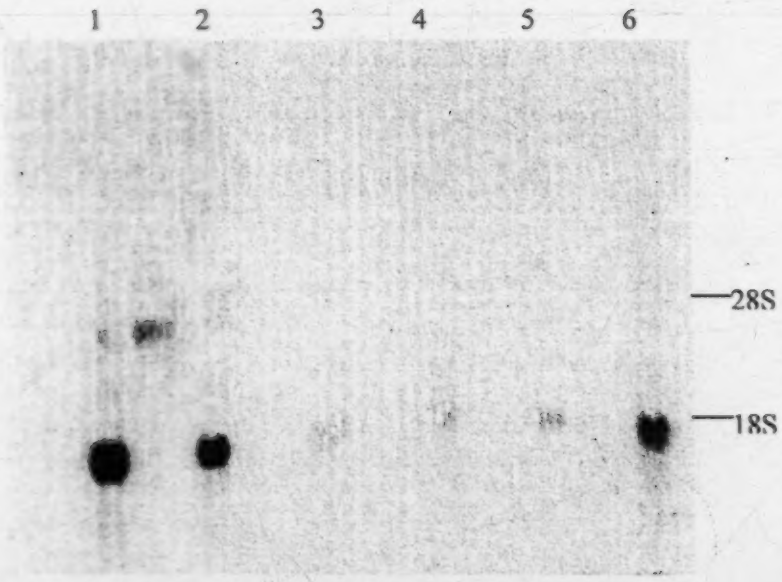
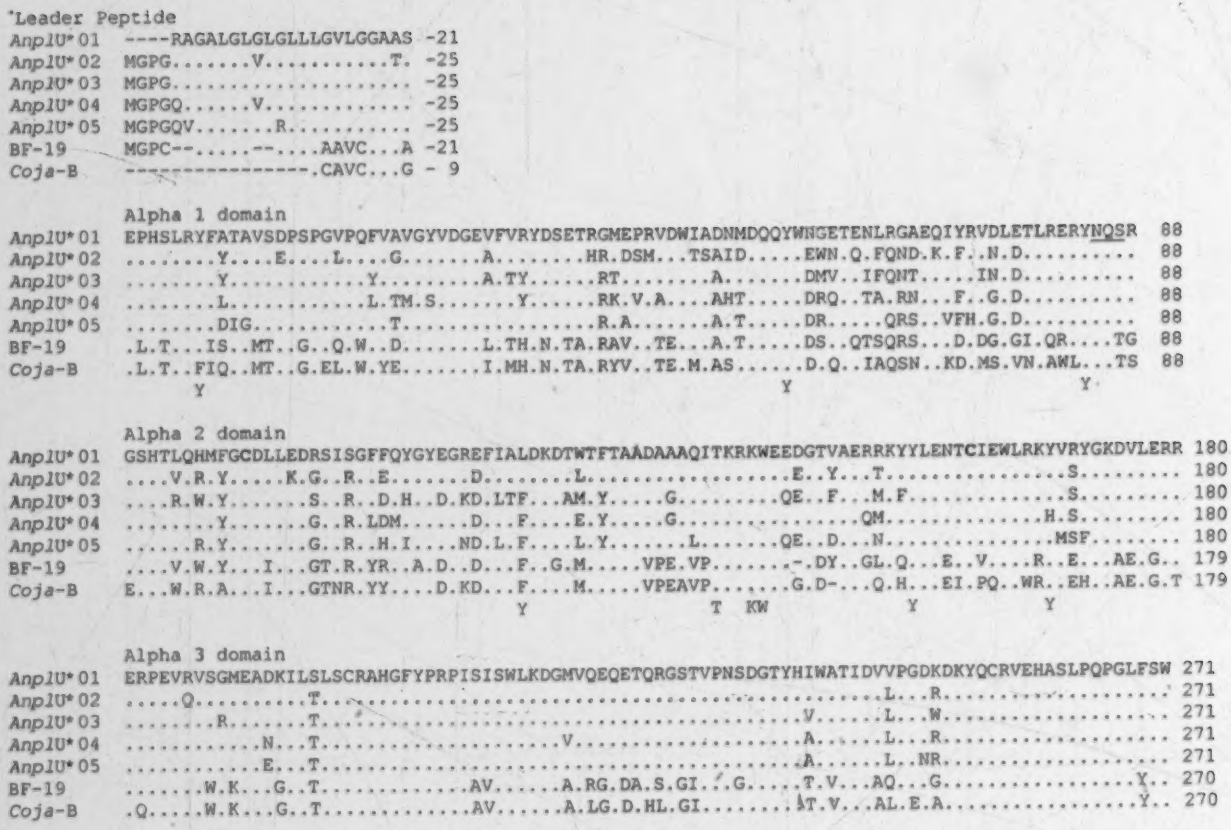


Figure 7: Northern blot analysis of *Anpl*U expression in different duck tissues. mRNA from the spleen of D26 and total RNA from various tissues of D132 were separated, transferred to a blot and hybridized with the MHC class I probe D26E4p. Positions of the 28S and 18S rRNA are labeled for reference. 1: spleen mRNA; 2: spleen, total RNA; 3: heart; 4: kidney; 5: testis; 6: liver.



* Figure 8: Alignment of the predicted amino acid sequence of the four MHC class I cDNA sequences (continued).

```
          Transmembrane domain
Anplu*01  E-PQSNLIPIVAGVSVAVVAV-IAALAGFAVWKSKQ 305
Anplu*02  .-.............A......-.............. 305
Anplu*03  .-.........V..A......-........I..... 305
Anplu*04  .-.............A......-.............. 305
Anplu*05  .-.........V..A......-.............. 305
BF-19     .P..P..V......A..I..IA.VVGV..IIYRRHA 306
Coja-B    .W..PSLV..IVR.VIPI.VIA.T.GI..TIYRRHA 306

          Cytoplasmic domains
           I           II          III
Anplu*01  GKKEKGYNLTP GSDGGSNSSNA GSNPSV  333
Anplu*02  ...G....VA. ..E........ ......  333
Anplu*03  ...G....VA. ..E........ ......  333
Anplu*04  ...GND..MA. ..........T ......  333
Anplu*05  ........VA. GMML                320
BF-19     ...G....IA. DRE...S..ST ....AI  334
Coja-B    ..K.....AA. SQ....SS.CT ...QTL  334
```

Figure 8: Alignment of the predicted amino acid sequence of the four MHC class I cDNA sequences to the previously submitted duck sequence *Anpl*U*01 (AF393511). The first amino acid of the α1 domain is considered as position 1. The chicken BF19 (Z54360) sequence, and quail Coja-B (AB005300) are included for comparison. Identity with *Anpl*U*01 is shown by a *dot*, and *dashes* indicate gaps that have been inserted to maximize the identity between sequences. Conserved amino acid residues at positions known to bind the peptide-terminal main chain atoms, are indicated below the alignment. The N-linked glycosylation site is underlined and the cysteine residues predicted to be involved in disulfide bonds, are shown in bold. The Y and S residues in the cytoplasmic domain that are involved in endocytosis of the molecule are underlined. Dots indicate identity with the sequence U*01. Dashes are included to maximize the alignment between sequences.

Exon boundaries were based on intron-exon data from PCR amplification of genomic DNA from D26 and confirmed by the typical intron-exon organization of other MHC class I molecules. In each sequence, we identified a leader peptide of 25 amino acids, an α1 domain of 88 amino acids, an α2 domain of 92 amino acids, and α3 domain of 91 amino acids, a transmembrane region of 34 amino acids, and a cytoplasmic region of 28 amino acids except for U*05 which had a cytoplasmic domain of only 15 amino acids. The percent amino acid identity of the complete sequences ranged from 84.4% to 85.9% (Table 3), which was lower than the range of percent nucleotide identity (88.5% to 91.1%). This difference is most apparent when the domains that form the peptide binding cleft, α1 and α2, are compared. The biggest contrast is between the α2 domains, where amino acid identity ranges from 65.9% to 79.6% while nucleotide identity is from 80.3% to 87.1%. This high level of amino acid sequence polymorphism may be the result of selection.

In comparison with the reported duck sequence U*01, we noted several amino acid differences were shared with all four of our sequences. We examined our four deduced amino acid sequences for separation into obvious pairs of alleles but they exhibited unique differences in each domain. It was not possible to assign these sequences to the same or separate loci by inspection. The majority of polymorphic positions lie within the α1 and 2 domains with 37/88 and 28/92 respectively, while only seven polymorphic positions are observed in α3. Many of the differences were not conservative changes. The differences between deduced amino acid sequences reflect alleles with very different peptide binding characteristics.

To determine whether all four sequences represent classical class I alleles, the sequences were examined for conservation of residues involved in peptide binding. There are highly conserved amino acid residues that anchor the peptide-terminal main-chain atoms (Y7, Y59, Y84, Y123, T143, K146, W147, Y159, Y171) that are generally conserved throughout evolution in classical class I molecules (Shum et al., 1999). Seven of the nine residues were conserved among the five duck cDNA sequences (Y7, $Y58^{Y59}$, Y84, $T140^{T143}$, $K143^{K146}$, $W144^{W147}$, $Y157^{Y159}$). Residue $Y168^{Y171}$ was conserved in U*02, U*03 and U*05 but not in U*04 which has H168 instead. The histidine residue 168 is seen in some classical sequences from human, mouse and rat (Shum et al., 1999). Y123 is not conserved in any of the aligned avian sequences but the leucine or phenylalanine at amino acid position 123, are seen in non-mammalian species (Shum et al., 1999). Other positions conserved among MHC class I molecules, as indicated by alignments of sequences from a variety of species (Kaufman et al., 1994), are the N-linked glycosylation site in α1 ($Q86^{Q87}$) and disulfide bonds in α2 ($C99^{C101}$ and $C162^{164}$). Also associated with classical class I molecules are a tyrosine in exon 6 (Y315 in *AnplU*) and a serine residue in exon 7 (S328) (Kaufman et al., 1994). It is known that the conserved tyrosine (Guild et al., 1983) and serine residue (Pober et al., 1978) in the cytoplasmic domains of MHC class I molecules, can be phosphorylated in the cell and may be involved in recycling of the class I molecules (Vega and Strominger, 1989).

A few insertions or deletions distinguish the duck sequences from the chicken and quail. *Coja*-B and BF19 shared an insertion of two amino acids in the transmembrane domain in comparison to the duck sequences: A294 in both and W274 and P274

respectively. Also, *Coja*-B and BF19 each had one deletion (duck amino acid positions 147 and 150 respectively) in α2 when compared to the duck sequences.

3.3 Amplification of MHC class I intron sequences from genomic DNA

To isolate the full genomic sequence of MHC class I, and identify the intron-exon organization, we amplified from the signal peptide (primer D26SP2) to the first cytoplasmic domain (D26CY1) using genomic DNA as template. The amplified product was cloned and sequenced and two sequences were identified which contained exons identical to alleles U*02 (9) and U*03 (18) (Table 4; Figure 9). The intron-exon junctions are all type 1. The introns of these two sequences were of identical length, shared 90% identity, and were quite short. Intron 1 was 150 bases in length, intron 2 was 665 bases, intron 3 was 94 bases, intron 4 was 83 bases, and intron 5 was 144 bases. These introns were similar in size to those of chicken MHC class I (Kroemer et al., 1990).

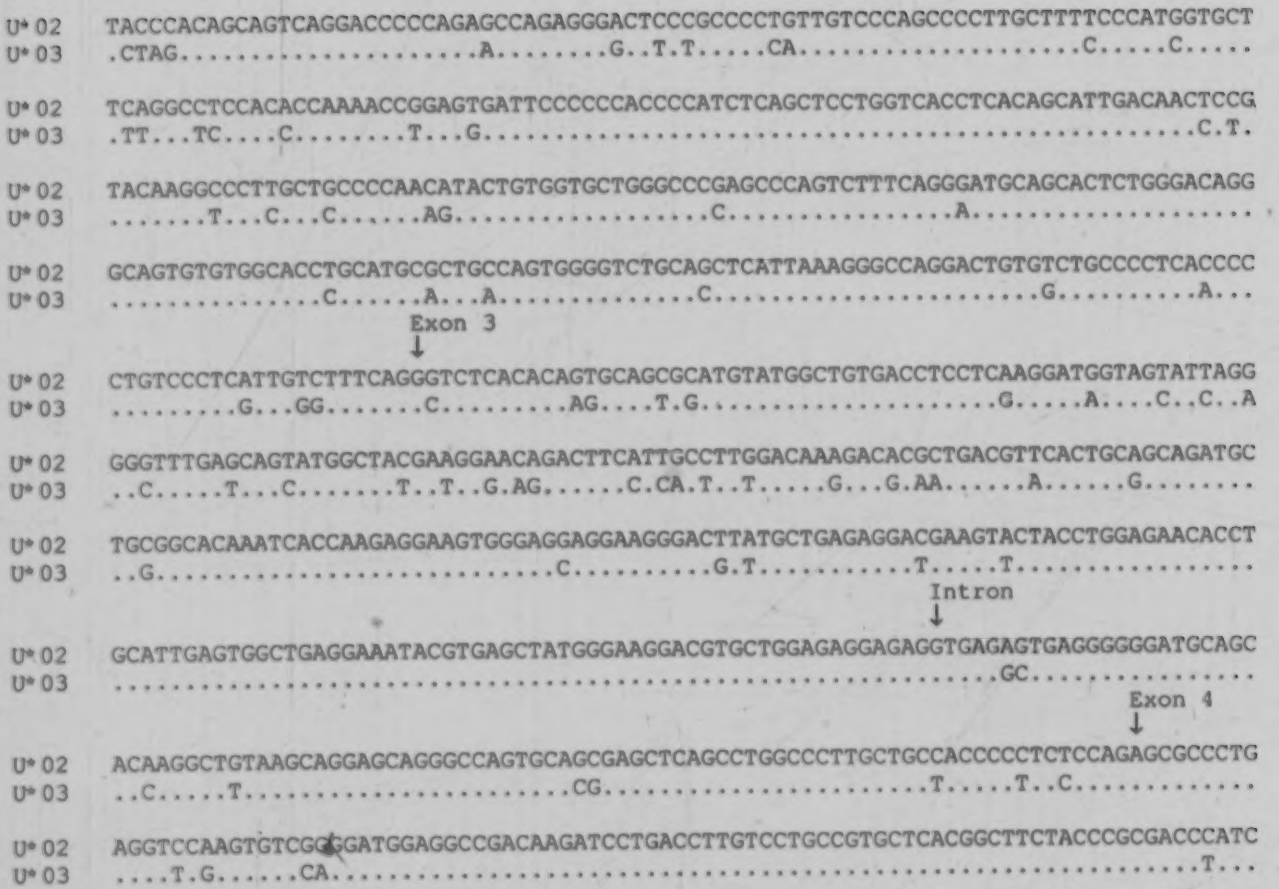
One sequence amplified from exon 1 to exon 6 shared identity with both U*02 and U*03 across certain regions (see appendix A, Figure A 1). This sequence, gU*12, is identical to the genomic sequence of U*03 in the first intron, differs from U*02 by two nucleotides in exon 2 but is identical to U*02 in the first 103 nucleotides of the second intron. The rest of gU*12 is identical to the genomic sequence of U*03. Though confirmation of this sequence is needed, gU*12 could be an allele at a distinct locus.

The forward primer D26SP2 also falsely primed in the reverse direction in intron 2 of some MHC class I sequences. This yielded two different products that did not span the entire targeted sequence (see Figure A 2 in Appendix A). One sequence was identical

Table 4: Summary of MHC class I sequences amplified by PCR from genomic DNA. Amplification across different introns was carried out with four sets of primers. Products were cloned and sequenced, and some sequences were different from the cDNA clones identified from the library. Numbers of clones of each sequence are identified in brackets.

	Sequences							
Amplification	U*02	U*03	U*04	U*05	One amino acid different from U*05	gU*12	gU*13	No identity with other alleles
Exon 1 to 6	(9)	(13)				(2)		
Exon 1 to Intron 2				(3)			(7)	
Exon 1 to 2	(1)	(2)	1-11(2)	1-9(2)	1-14(2)		1-10(2)	1-8(1)
Exon 2 to 3	(2)	(2)	2-4(3)	2-8(6)	2-1(6)			
Exon 3 to 4	(2)	(7)	4-5(1)	4-6(4)				3-12 (5) 3-14 (1) 3-1 (1) 4-1 (1)

* Figure 9: Nucleotide alignment of the MHC class I genomic PCR fragments of U*02 and U*03 (continued).



* Figure 9: Nucleotide alignment of the MHC class I genomic PCR fragments of U*02 and U*03 (continued).

Figure 9: Nucleotide alignment of the MHC class I genomic PCR fragments amplified from exon 1 to exon 6. Primers D26SP2 and D26CY1 were used to amplify MHC class I products from D26. The PCR products were cloned and sequenced. Sequences containing coding region sequences identical to the cDNA clones U*02 and U*03 were identified and aligned. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02.

to U*05 in exon 2 (three clones). Another sequence from this amplification, designated gU*13 (seven clones) was identical to a sequence amplified with the primers UTF2 and E2R1 (see Table 4 and below). The putative exon 2 translation of gU*13 contained the conserved amino acids necessary to bind peptides.

Fragments spanning introns 1, 2, and 3 were also amplified separately to determine the possible intronic sequences of other MHC class I alleles. The amplification from the 5' untranslated region to exon 2 (primers 5'UTF2 and D26E2R1), yielded sequences that had coding region sequences similar to all four alleles identified in the cDNA library. One sequence matched U*02, one sequence matched U*03 (two clones), sequence 1-11 (Figure A 3 in Appendix A) matched U*04 (two clones) and one sequence (1-9, 2 clones) matched U*05 in the coding region. One sequence (1-14, 2 clones) differed from U*05 by two nucleotide mismatches within exon 2, leading to the amino acid change A44E. This change, relative to U*05, is not a PCR generated artifact. The same change in exon 2 relative to U*05, was present in sequence 2-1 (6 clones) generated by a separate PCR that amplified across intron 2 using the primers D26E2F1 and D26E3R1 (see Figure A 4). Sequence 1-10 (two clones) (Figure A 3) was identical to the sequence gU*13 that was amplified from exon 1 to intron 2. There was also one sequence (1-8, one clone) that did not match any other sequences identified.

The amplification across the intron 2 also yielded sequence 2-8 (6 clones) (Figure A 4) identical to U*05 in the coding region. From the same PCR amplification, one sequence (2 clones) matched U*02, another sequence (2 clones) matched U*03, and sequence 2-4 (3 clones) was found to match U*04.

The genomic amplification across the third intron was done with two pairs of primers: D26E3F1 with D26E4R2, and D26E3F2 with D26E4R2. There were sequences that matched U*02 (2 clones) and U*03 (7 clones). Sequence 4-5 (1) (Figure A 5) was identical to U*04 in the coding region and sequence 4-6 (4) was identical to U*05. Sequence 3-12 (5) and three other sequences, 3-1 (1), 3-14 (1) and 4-1 (1), did not match any sequences previously identified. It is possible the three sequences for which only one clone was identified, were PCR generated artifacts since they have so few differences between them, and no other clones identical to them were sequenced.

3.4 Identification of the TAP2 gene

We anticipated that the biased representation of duck MHC class I alleles could be due to close association of one locus with TAP genes. To examine the organization of the MHC region of ducks we cloned the duck TAP genes. A probe (D26T2WAB) for TAP2 was amplified from genomic DNA from the Walker A sequence in exon 6 to the Walker B sequence in exon 8, conserved sequences in the ATP-binding cassette of TAP2. To identify expressed TAP2 sequences, the spleen cDNA library prepared from D26 was screened. Five plaques hybridized with the probe and were excised into plasmids for sequencing. The clones contained TAP2 sequences with 83.2% homology with chicken (Figure 10), but all were incomplete at the 5'end, starting in exon 6 or 7. No full-length cDNA was obtained.

To determine the complete TAP2 sequence, we designed a primer (CHT2E1F1) in the 5' end based on a sequence conserved between chicken and quail TAP2. We used a PCR approach to assemble a complete TAP2 sequence from the genomic DNA of D26. Overlapping genomic TAP2 PCR products were cloned and sequenced. Cloned products,

* Figure 10: Nucleotide alignment of duck and chicken TAP2 cDNA's (continued).

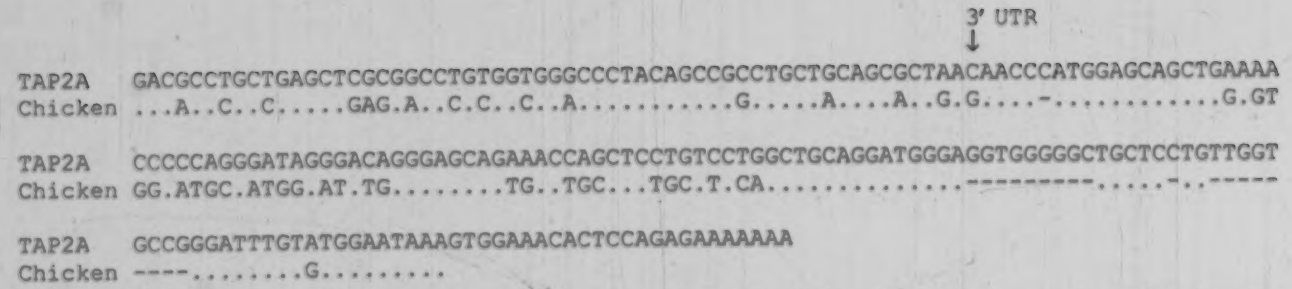


Figure 10: Alignment of duck TAP2 cDNA with chicken TAP2 (Genbank Accession number AL023516). Arrows mark the beginning of exons. Dots show identity with the TAP2A sequence and dashes are included to maximize the identity between the sequences.

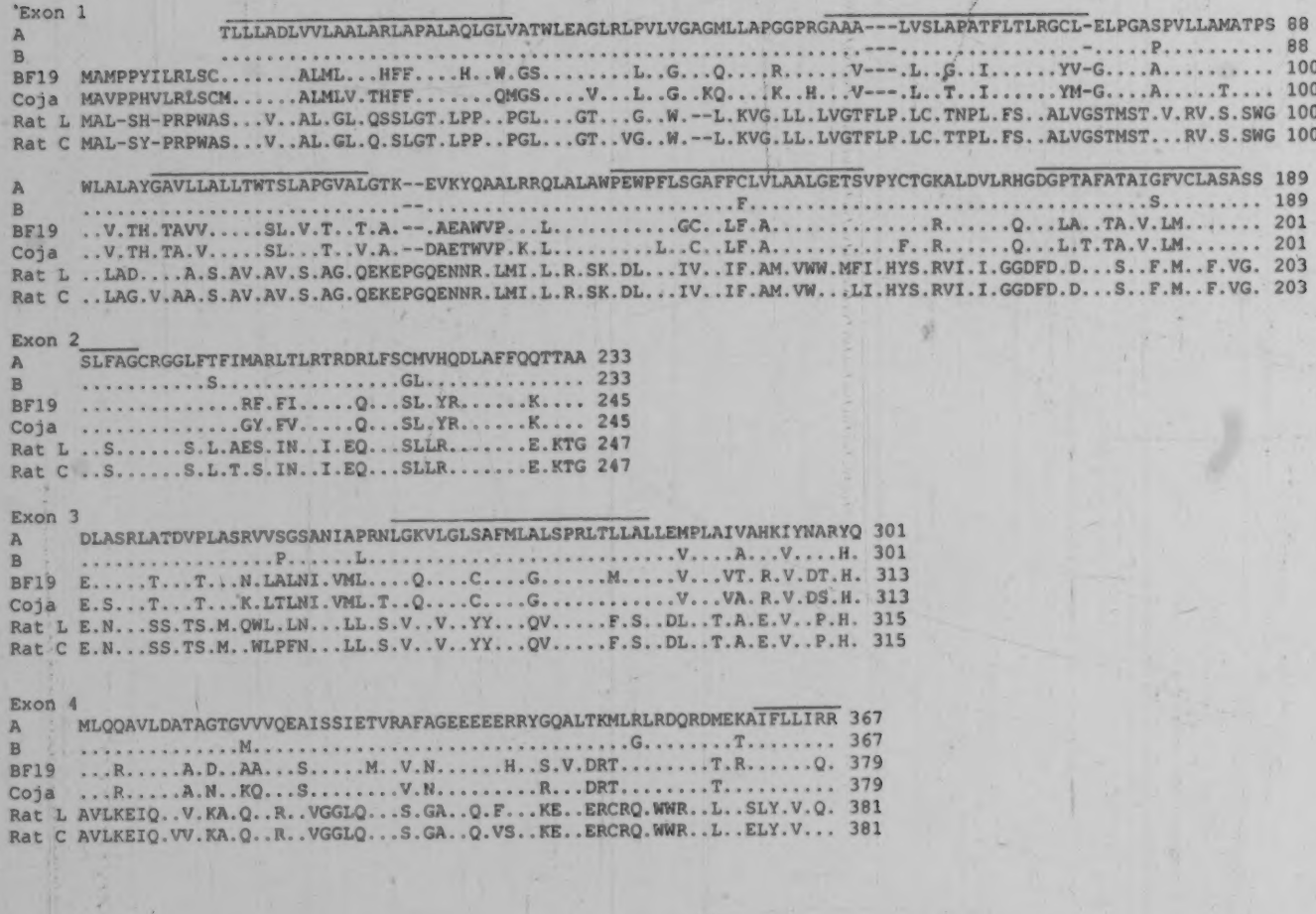
identical in the overlapping regions, were assembled and two alleles (A and B) were identified. The assembled contigs of TAP2 A and B were 3010 nucleotides in length (see Figure A 6). The intron/exon boundaries were determined from the intron/exon splice sites and then confirmed by comparison of the sequences with TAP2 sequences from chicken and quail. TAP2 A and B share 96.6% nucleotide identity when comparing both the introns and exons of the contigs. Duck TAP2A shares 73.9 % and 70.6% nucleotide identity with chicken and quail respectively, while TAP2B shares 75.6% and 74.3% nucleotide identity with chicken and quail. If only the nucleotides of the exons are compared, TAP2 A and B share 97.3 % identity, TAP2A is 80.4% and 79.3 % identical to chicken and quail, and TAP2B is 81.0% and 79.9% identical.

3.5 Reverse transcriptase PCR amplification of TAP2

Because only one allele was identified by screening the cDNA library, we amplified a fragment of TAP2 from D26 spleen mRNA by RT-PCR to determine if both TAP2 alleles were expressed. We used primers common to both alleles in exon 4 (primer DT2E4F) and exon 8 (DT2E8R1). The amplified product was cloned, sequenced, and sequences corresponding to TAP alleles A and B were identified.

3.6 TAP2 amino acid sequences

The amino acid sequence of TAP2A and TAP2B was deduced from the exons of the genomic contigs and aligned with the chicken, quail and rat TAP2 (Figure 11). Alleles A and B shared 96.5% amino acid identity over 689 amino acids. The sequences contained the highly conserved Walker A and Walker B motifs (Walker et al., 1982) that bind ATP, and the signature motif (Loo et al., 2002) thought to bind to the opposing Walker A sequence in the nucleotide binding domain of the adjacent TAP1. TAP2A and



* Figure 11: Amino acid alignment of duck TAP2 with chicken, quail and rat TAP2 (continued).

```
Exon 5
A     ALQLAVQALVLYCGHQQLHEGTLTAGGLIAFILYQNNAGSSVQ 410
B     ..................R.........V.............. 410
BF19  V.........................V......TK...C.. 422
Coja  ...................D......S.V......T....C.. 422
Rat L VMA.GM.V.I.NV.V..ILA.EV.R...LS.L...EEV.HH.. 424
Rat C VMA.GM.V.I.N..V..ILA.EV.R...LS.L...EEV.HH.R 424

Exon 6
A     ELAYAYGNLLSNVAAAHKIFDYLDREPAVGTGGTLEPATLQGHVTFQHVSFAYPTRPEHLVLQDVSFELRPGEVTALAGLNGSGKSTCAGLLQRFYEPTAGEVLLDGVPLRDYKHRYLHRQ 531
B     A......D.........................S.....................................................R..... 531
BF19  A...S..D....A...C.V....NW.R...A...YV.TR.R.....HR..........R......T.....................VA..E.....G.............E..... 545
Coja  A...C..D....A...C.V.......R.......YV.TK....I..HR.................T.....................VA..E.....K...........Q..E.C..... 545
Rat L N.V.M..DM....G..E.V.S....R.NLPNP...A.PR.E.R.E..D...S..S...KP...GLT.T.H..K....V.P.......V.A...NL.Q..G.QL....E..VQ.D.H..... 547
Rat C N.V.M..DM....G..E.V.S....R.NLPKP...A.PRVE.R.E..D...S..S...KP...GLT.T.H..K....V.P.......V.A...NL.Q..G.K.....E..VQ.D.H..... 547

Exon 7
A     VALVGQEPVLFSGSIRDNIAYGLEDCREEEIIAAAEAAGALDFISALDQGFDTD 585
B     ..................................................... 585
BF19  ......................M...E........R.....G.....E...G.. 597
Coja  ......................M...E........R.....G.....E...G.. 597
Rat L .V............VK.......R..EDAQVM...Q..C.D...GEMTN.IN.E 599
Rat C .V............VK.......R..EDAQVM...Q..C.D...GEMTN.IN.E 599

Exon 8
A     VGERGDQLSAGQKQRVAIARALVRCPTVLILDEATSALDGDGDVP 630
B     ..............................................A 630
BF19  .....G.........I........H..................S.AM 642
Coja  .....G.........I........H..................M 642
Rat L I....S..AV.....L........N.R............AECEQA 644
Rat C I..K.S..AV.....L........N.R............AECEQA 644

Exon 9
A     LQQWVQSGGDRTVLLITHHPRMLEKADCVVVLERGTVAEMGTPAELAACGGPYSRLLQR 689
B     ..................W......RI....H............T..... 689
BF19  .....RN...........Q........RI....H............RTR.........H 701
Coja  .....RN.V.......S.Q........RV....H.M.......T..RTR.........H 701
Rat L ..T.-R.QE...M.V.A.RLHTVQN..Q.L..KQ.QLV.HD---Q..DEQDV.AH.V.Q 699
Rat C ..T.-R.QE...M.V.A.RLHTVQN..Q.L..KQ.QLV.HD---Q..DEQDV.AH.V.Q 699
```

Figure 11: The predicted amino acid sequence of duck TAP2 alleles A and B. Alignment includes chicken TAP2 (AL023516), quail TAP2 (AB007195) and two rat TAP2 alleles (Rat $Tap2^{L}$-X75305; Rat $Tap2^{C}$-X75306). Identity with duck TAP2 A is shown by a *dot*, and *dashes* have been included to maximize the identity of the alignments. The predicted locations of the Walker A and Walker B sequences are underlined. Predicted transmembrane helices are overlined. Dots show identity with the TAP2A sequence and dashes are included to maximize the identity between the sequences.

TAP2B shared 77.1% or 77.8% amino acid identity with chicken respectively and 76.6% and 77.1% identity with quail. Chicken and quail amino acid sequences were 90.3% identical.

Out of a total of 55 nucleotide differences, there were 26 amino acid differences between TAP2 allele A and allele B. Eighteen of the amino acid differences between the alleles are non-conservative changes in terms of charge and size. If we compared the number of non-synonymous substitutions to synonymous, the ratio was less than one over the whole gene, suggesting that there was no selection for polymorphism in TAP2. However, these amino acid differences were not distributed evenly throughout the sequence. There were two non-conservative amino acid differences in exon 2 (C218G and M219L), four in exon 3 (S251P, P258L, M286V, and Y300H), three in exon 4 (V316M, R350GG, and K359T), three in exon 6 (E411A, L445S, and P630A), and three in exon 9 (R651W, C658R, and A677T). Potentially functional differences in the two rat TAP2 sequences were clustered in the same regions. The function of the TAP2 molecule is to bind and transport peptides into the ER. An amino acid change in the area that binds peptides can alter the peptide binding specificity of the molecule. We believe some of the functional amino acid differences were clustered in areas of TAP2 responsible for the peptide binding specificity of the molecule. To determine the approximate positions in the molecule where the amino acid differences would occur, we generated a model of the protein structure for duck TAP2 (Figure 12) identifying the position and orientation of the putative transmembrane (TM) helices. There are many conflicting opinions on the number and location of TM helices in TAP2 (Momburg et al., 1996; Nijenhuis and Hammerling, 1996; Velarde et al., 2001; Vos et al., 1999). After analyzing the output of

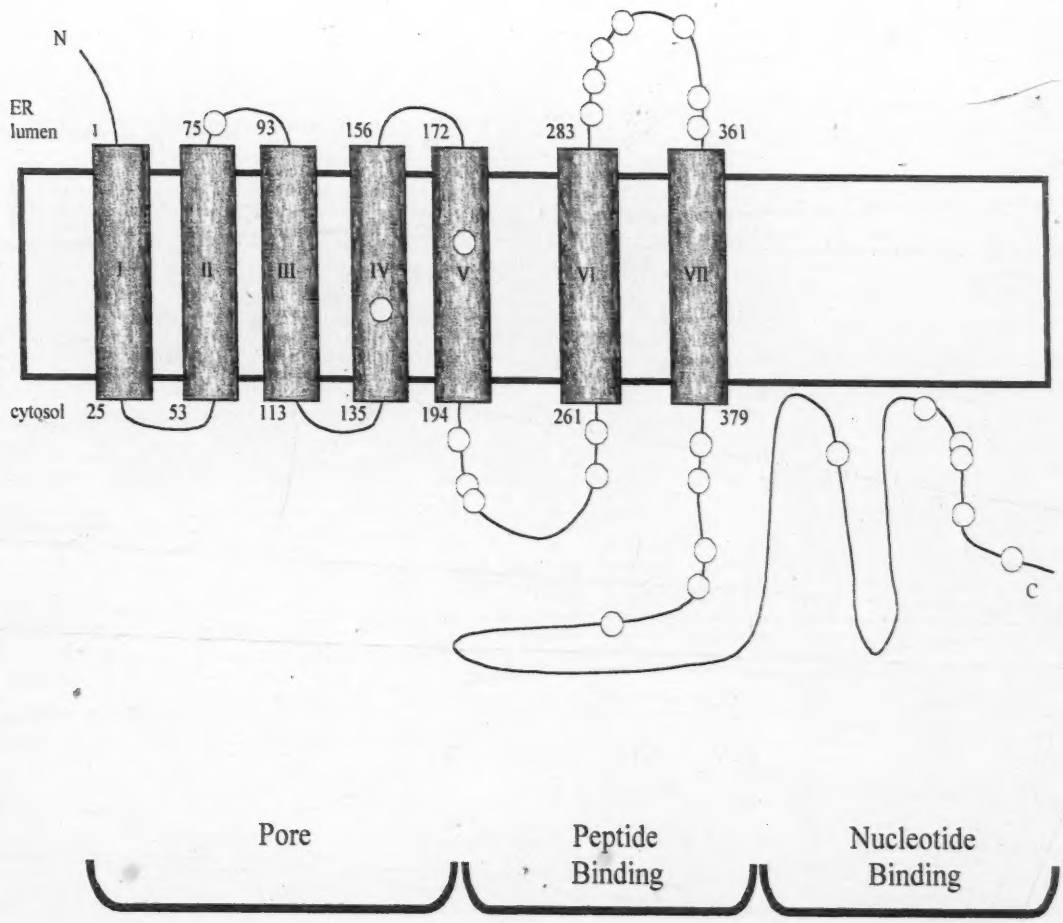


Figure 12: Model of the duck TAP2 protein. The amino acid differences between duck TAP2A and TAP2B occured in positions involved in binding and transporting peptides. Amino acid residues that form the transmembrane helices are numbered according to the duck TAP2 alignment. Circles represent the approximate position of the 26 amino acid differences between the two duck TAP2 alleles.

multiple programs that predict hydrophobic regions, we predicted that there are 7 possible TM helices. Our model correlated with previous analyses (Velarde et al., 2001; Vos et al., 2000; Vos et al., 1999), which placed the N-terminus of the TAP2 protein in the ER lumen. These studies also predicted the first five TM segments make up the pore region, and that peptide binding and specificity is determined by loops of the protein in the cytosol after the fifth TM helix. For duck TAP2, although there are differences throughout the molecule, amino acid changes cluster in regions likely important for peptide binding (Nijenhuis and Hammerling, 1996; Vos et al., 1999).

3.7 Amplification and sequencing of the intergenic region between the TAP2 locus and the dominantly expressed MHC class I

In the chicken, the major MHC class I locus is downstream of the TAP2 locus. The loci are in opposite transcriptional orientation, and are separated by only 1.8 kb (Kaufman et al., 1999b). To determine if the duck also had an MHC class I locus adjacent to TAP2, we used a primer from the 3' end of *Anpl*U (D26TMF), common to the cDNA sequences identified in D26, and a primer from the 3' end of the TAP2 locus (DT2E7F) to PCR amplify between the genes. We cloned the amplified product from two different ducks, sequenced eight clones from each by primer walking, and identified two distinctly different sequences. In both D26 and D95, the D26 TAP2B allele and the MHC class I allele U*02 were at opposite ends of one clone. The exons and 3'untranslated regions of the other cloned sequence was identical to TAP2A and U*03 (see Figure 13 and Appendix A, Figure A 6 and Figure A 7). Isolating the insert from the vector by restriction digest with *Eco*RI, showed that the haplotype bearing TAP2 A

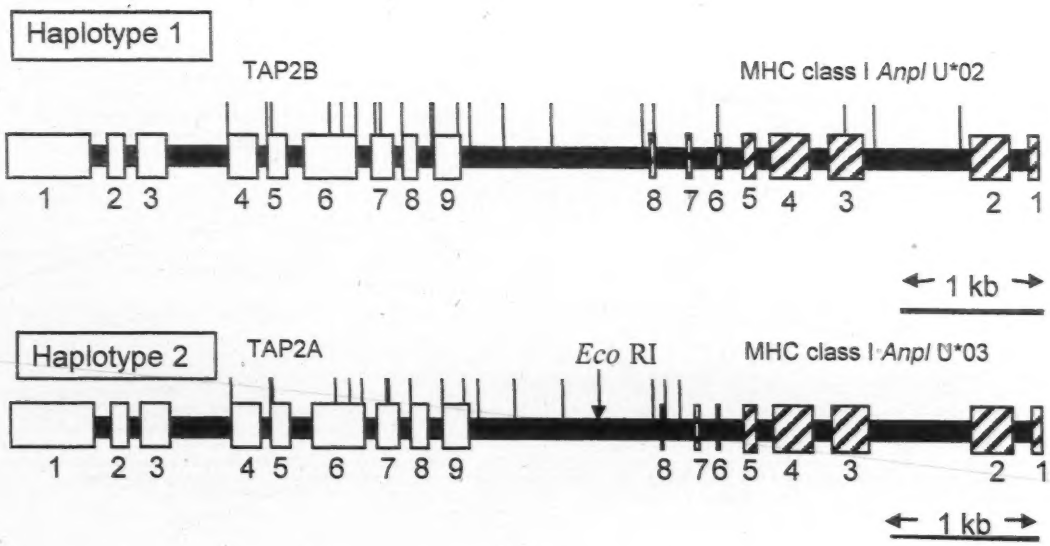

Haplotype 1
TAP2B
MHC class I Anpl U*02
1 2 3 4 5 6 7 8 9 8 7 6 5 4 3 2 1
1 kb
Haplotype 2
TAP2A
Eco RI
MHC class I Anpl U*03
1 2 3 4 5 6 7 8 9 8 7 6 5 4 3 2 1
1 kb

Figure 13: Genomic organization of duck TAP2 and the putative major MHC class I locus. The lengths of the exons and introns are based on PCR amplification. The sites where the enzyme *Pst*I cuts in each haplotype are shown by a line. The only *Eco*RI cut site identified was in haplotype 2 (arrow). Haplotype 1: TAP2B and *Anpl*U*02; haplotype 2: TAP2A and *Anpl*U*03.

contained an *Eco*RI site within the intergenic region near the 3' end of *Anpl*U*03 whereas the other haplotype did not.

3.8 Southern hybridization

To estimate the number of MHC class I loci, we performed Southern blot hybridization. We chose the enzyme *Pst* I because it revealed restriction fragment length polymorphisms (RFLP) between different quail (Shiina et al., 1999b). Also, *Pst* I did not cut within the exon 4 probe used. *Pst* I digested duck genomic DNA was hybridized with the MHC class I exon 4 probe D26E4p under low and high stringency conditions (Figure 14). Of nine individuals analyzed, seven ducks had four hybridizing fragments. Ducks 26, 95, 132 and 213 (lanes 1, 2, 5 and 7) had fragments approximately 2.4 kb, 2.2 kb, 1.3 kb and 0.8 kb in size. In three ducks, 105, 129 and 138 (lanes 3, 4 and 6), the 2.4 kb and 2.2 kb fragments were slightly larger (approximately 2.5 kb and 2.3 kb). Only the 1.25 kb and 0.8 kb fragments hybridized in duck 180 (lane 8), and only the 0.8 kb band hybridized in D64 (lane 9). The level of stringency used did not affect the hybridization pattern seen in the ducks. However, high stringency decreased the number of hybridizing bands in the two chicken samples. Five bands in one chicken sample and nine bands in the other sample were visible at low stringency.

Hybridization of MHC class I with duck genomic DNA shows there are no more hybridizing bands in the duck than in the chicken. Four hybridizing bands in seven of the ducks suggested there could be at least four MHC class I loci but amplification of genomic sequences suggests there are more. Analysis of the genomic U*02 sequence found *Pst* 1 sites in exon 3 and exon 6 (Figure 13) create a 0.8 kb fragment, matching the approximately 0.8 kb southern fragment. The first *Pst* I site in U*03 is 2.4 kb from the

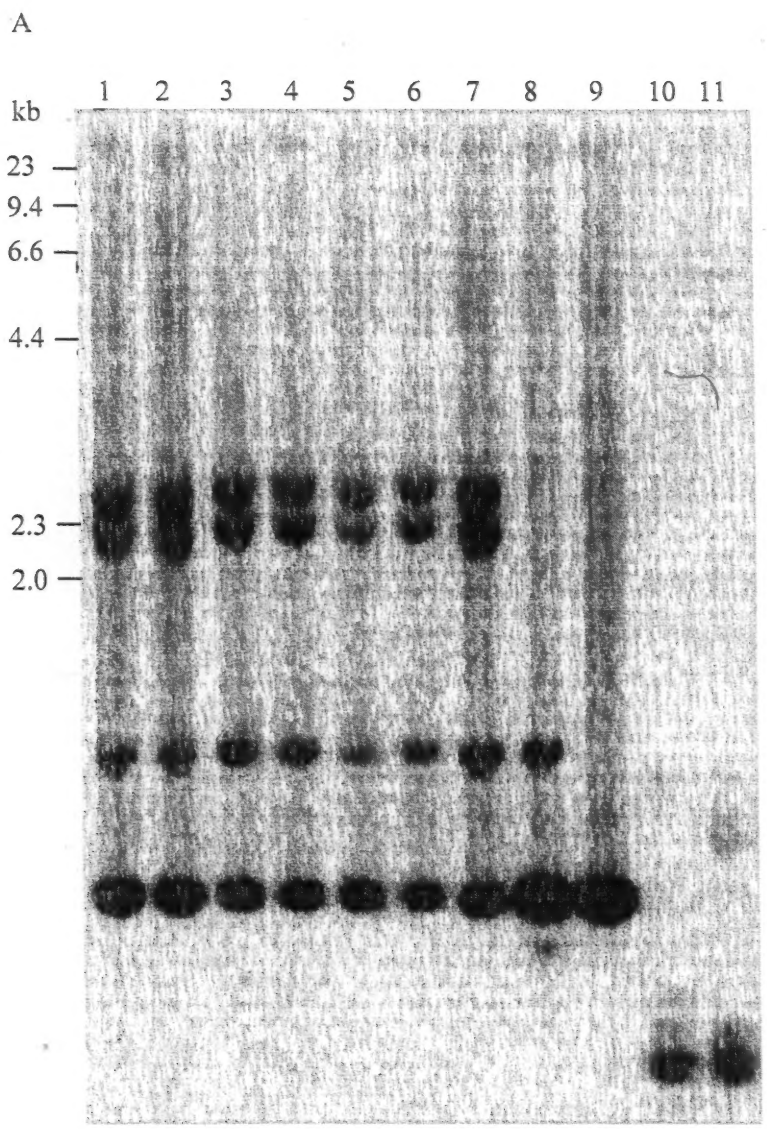


*Figure 14: MHC class I hybridization in ducks and chickens (continued).

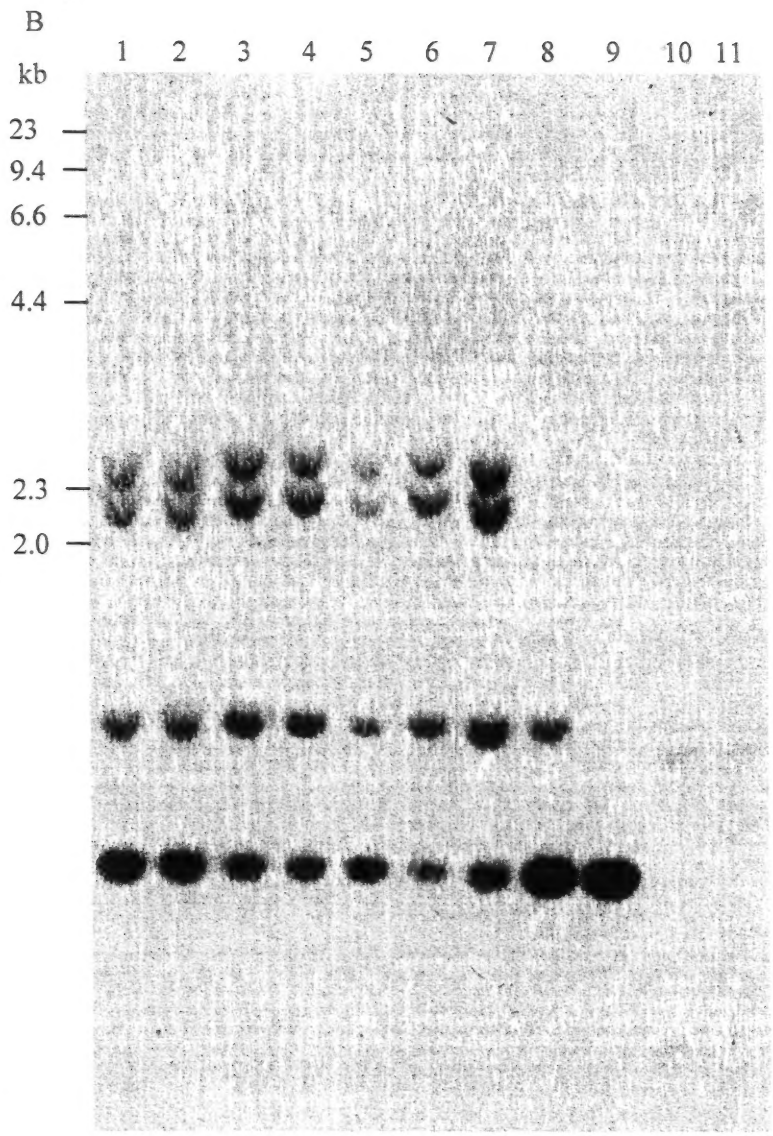


Figure 14: MHC class I polymorphism in ducks and chickens. *Pst*I digested genomic DNA of nine ducks (lanes 1-9: D26, D95, D105, D129, D132, D138, D 213, D180, D64) and two chickens (lanes 10 and 11) were hybridized to the MHC class I probe D26E4p under low stringency (A) and high stringency (B) conditions. Sizes of HindIII-digested lambda DNA are shown.

start codon of this allele. Though we do not know where the *Pst* I site is upstream of the gene, this fragment could correspond to the band observed that is approximately 2.4 kb.

We also performed southern blot hybridization to observe similarities between MHC class I and TAP2 fragment length polymorphisms among the ducks (Figure 15). Duck genomic DNA digested with *Eco* RI was hybridized sequentially with the probes for TAP2 (tap2WABp) and MHC class I (D26E4p). Two patterns were observed among the five ducks analyzed. D26, 95, and 132 have two TAP2 hybridizing fragments, 7.1 kb and 6.7 kb, while D105 and D129 only have one hybridizing fragment at 6.7 kb. Individuals that have a similar pattern of hybridization for TAP2 have a similar hybridization pattern for MHC class I when the DNA is digested with either *Pst*I (Figure 14) or *Eco* RI (Figure 15A).

We know that TAP2 and MHC class I are linked based on sequencing data from D26. Haplotype 2 (TAP2A/U*03) in D26 contains an *Eco*RI restriction site between the two genes. This restriction site does not exist in the other haplotype (TAP2B/U*02). Only the larger TAP2 fragment (Figure 15) cross-hybridizes with the MHC class I probe. Looking at the banding pattern seen in D26 for TAP2, the larger fragment could be from hybridization with the TAP2B allele, and the smaller band observed would be from hybridization with the TAP2A allele that is separated from the class I locus by *Eco*RI digestion (Figure 13). Similar haplotypes occur in D95 and may be found in D132, whereas different haplotypes may exist in D105 and D129. MHC class I and TAP2 also hybridized to restriction fragments of identical size in genomic DNA digested with *Hin*dIII (Figure 16). Two ducks (D26 and D132) were compared. The same blot was sequentially hybridized with a TAP2 probe and an MHC class I probe. Each probe

hybridized with only one fragment, which was approximately 23 kb in size. The hybridizing band in D26 is slightly larger than the one for D132.

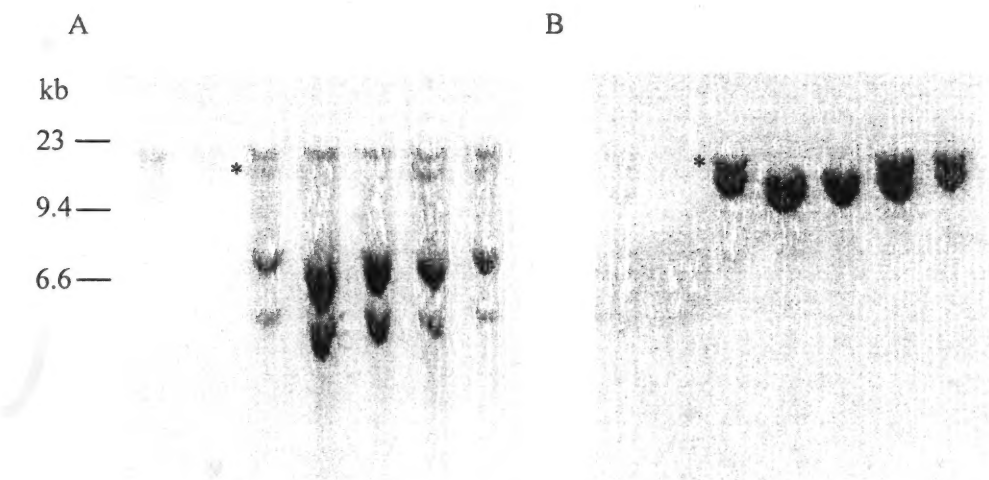


Figure 15: Comparison of MHC class I and TAP2 hybridization patterns in five ducks. Genomic DNA from two chickens (lanes 1 and 2) and five ducks (D95, 105, 129, 132 and 26 in lanes 3-7 respectively) was digested with the enzyme *Eco*RI and hybridized sequentially to MHC class I (A) and TAP2 (B) under high stringency conditions. An asterisk (*) indicates hybridizing bands of the same size between the two blots. Sizes of the *Hin*dIII digested lambda DNA are shown.

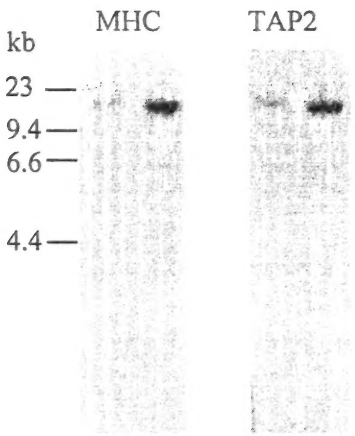


Figure 16: MHC class I and TAP2 hybridize to restriction fragments of identical size. DNA from related ducks, D26 (first lane) and D132 (second lane), was digested with *Hin*dIII, and the same blot was sequentially hybridized with MHC class I and TAP2 under high stringency conditions. Sizes of the *Hin*dIII digested lambda DNA are shown.

4 DISCUSSION

We have examined the expression and linkage of MHC class I genes in the duck. Southern blot analysis and amplification of genomic sequences suggested that ducks have at least four loci encoding MHC class I. However, one sequence was over-represented in the unamplified cDNA library. We cloned the TAP2 gene from the cDNA library and isolated two alleles of the TAP2 gene by PCR amplification from genomic DNA. RT-PCR amplification of the alleles showed that they are both expressed. PCR amplification from the class I gene to TAP2 allowed us to determine that the dominantly expressed MHC class I locus is tightly linked to the polymorphic TAP2 locus.

4.1 MHC class I expression in the duck

We found from the unamplified duck spleen cDNA library that out of four expressed MHC class I transcripts, two sequences, U*02 and U*03 accounted for the majority of MHC class I clones isolated. All four expressed sequences have the characteristic residues that anchor the main chain terminal atoms, and are conserved in classical class I molecules (Shum et al., 1999). These alleles also differ significantly within the peptide-binding region, and if expressed, could contribute to functional diversity by presenting a different profile of peptides. Since this library had not been through any rounds of amplification after phage packaging, the number of clones isolated reflects the expression level of the alleles in mRNA.

The dominant expression of a single class I locus has, thus far, only been observed in non-mammalian vertebrates. Among the teleost fish such as salmonids (Aoyagi et al., 2002: Shum, 2001; Grimholt et al., 1993), the guppy (Sato et al., 1995), catfish (Antao et al., 1999) and carp (van Erp et al., 1996), there is evidence of only one

expressed MHC class I locus despite the existence of more than one locus. The amphibian *Xenopus* is also known to have one expressed classical class I locus (Sammut et al., 2002) with a number of non-classical class I loci (Flajnik et al., 1993). The chicken has been the only avian species for which MHC class I expression has been extensively characterized and the chicken shows dominant expression of alleles from one locus (Kaufman et al., 1995).

Like the chicken, we have evidence that the dominantly expressed duck alleles U*02 and U*03 are alleles at a single locus. The chicken has a second classical MHC class I locus, called the minor locus, linked to the dominant locus. This minor locus is disrupted so it is not expressed, or expressed at a low level (Kaufman et al., 1999a). There are also two class I loci outside of the chicken MHC, in the Rfp-Y region (Miller et al., 1994). Only one of these loci is expressed and shows characteristics of being a non-classical class I locus (Afanassieff et al., 2001). We do not know the location of the loci encoding the less abundant U*04 and U*05 sequences. If the duck contains a region similar to the chicken Rfp-Y, unlinked from the MHC proper, it is conceivable that some of the expressed MHC class I loci reside there. The expression and organization of U*04 leads us to believe it could function as a classical class I molecule. However, the complete identity of U*05 is unknown. Only one complete cDNA clone of U*05 was found, and it did not possess a complete cytoplasmic domain, or an obvious polyadenylation signal. We also do not know how or why ducks limit the expression of these sequences, although we can speculate based on the closely related chicken.

4.2 Polymorphic TAP2 linked to dominant MHC class I locus

Amplification with primers from MHC class I and TAP2 showed the major MHC class I locus of ducks was tightly linked to TAP2. The 3' ends of these genes are separated by 1.3kb. The genomic organization of the MHC in both chicken and quail has shown tight linkage between these genes. Other non-mammalian vertebrates such as fish and amphibians as well as some mammals, such as rats, exhibit linkage between the antigen processing and presentation loci. This tight physical linkage may result in co-evolution between these genes.

The physical linkage of MHC class I with TAP may limit the usefulness of expressing additional loci. A particular TAP2 allele may evolve to function with a particular MHC class I allele to transport and present peptides with a certain motif. Other class I alleles in the haplotype must have similar peptide specificity to ensure loading. If they are not able to load peptides for presentation to the T cells, there can be no selection for retaining their function. Over time these alleles may accumulate mutations that further decrease their function. This scenario is proposed to have occurred in the chicken. In different chicken haplotypes, part of the promoter region of MHC class I alleles at the minor locus are either distorted or deleted, disrupting the expression of this locus (Kaufman et al., 1999a).

The peptide specificity of a TAP2 molecule determines which peptides are available to bind to and be presented on MHC class I. In the rat, polymorphism of TAP2 has been shown to lead to functionally different alleles that preferentially bind peptides with certain motifs. Rat TAP2A alleles (alleles $TAP2^a$ and $TAP2^l$) are considered permissive and bind almost any peptide, while TAP2B alleles (alleles $TAP2^c$ and $TAP2^u$)

only transport peptides with a hydrophobic C-terminal amino acid (Joly et al., 1996; Powis et al., 1996). The rat MHC class I allele RT1.A^a prefers to bind peptides with a C-terminal arginine (Powis et al., 1996). Polymorphic residues close to the ER membrane in the cytosolic loops after TM V are important in the observed differential binding of peptides (Deverson et al., 1998; Momburg et al., 1996). The *cim* effect, or class I modulation, was observed in a rat homozygous for RT1.A^a and TAP2^u. TAP2^u lacked the ability to transport peptides carrying the positively charged arginine at the C-terminus into the ER, limiting the peptide binding, and ultimately surface expression, of RT1.A^a.

Given the close linkage of the dominantly expressed duck MHC class I locus to TAP2, we predicted that TAP2 would also be polymorphic. Indeed we identified two alleles of TAP2 by amplification from genomic DNA. There were 26 amino acid differences between the two alleles identified in D26 (Figure 11). Several of the amino acid differences between these alleles were not conservative changes, and they appeared to be clustered in distinct regions of the protein, as is seen between the different rat TAP2 alleles.

We predicted the location of TM segments in the duck TAP2 sequence (Figure 12). Our analysis predicted seven TM segments, and placed the N-terminus of the TAP2 protein in the ER lumen, consistent with previous analyses (Velarde et al., 2001; Vos et al., 2000; Vos et al., 1999). The first five TM segments make up the pore region, and peptide binding and specificity is determined by loops of the protein in the cytosol after the fifth TM helix. Significant amino acid differences between the duck TAP2 alleles clustered to the cytosolic loop between TM helices V and VI, and near the ER membrane after TM helix VII. These are areas where amino acid differences between the rat TAP2

alleles are implicated in altering the peptide binding specificity of the transporter (Deverson et al., 1998; Momburg et al., 1996; Nijenhuis and Hammerling, 1996). Therefore the amino acid differences in the duck are located in regions that could influence peptide binding specificity. Functional studies are needed to determine if these amino acid differences do alter the transport of peptides.

4.3 The MHC of the duck

Southern blot hybridization with MHC class I suggested that ducks have at least four MHC class I loci. From PCR amplifications of genomic DNA, we have evidence of nine alleles, defined by unique exon 4 sequences (Table 2; Figure 5; Table 4). Only one, gU*10N, could not possibly be expressed. U*02 and U*03 are alleles at one locus, linked to the TAP2 locus. It is not known where in the genome the other alleles reside, either in the MHC or in a region analogous to the chicken Rfp-Y. A sequence very similar to U*05 was found by genomic amplification and likely is an allele at a locus with U*05. The locus designation of the other sequences (U*04, gU*06, gU*07, gU*10N and gU*12) are not known.

Partial evidence from several species of birds suggests that the minimal MHC of chickens is not typical of all birds. In a recent review (Kulski et al., 2002), the organization of the quail MHC region was presented. Quail have seven MHC class I loci in the MHC (Kulski et al., 2002), and express transcripts from multiple loci (Shiina et al., 1999a). Our preliminary analysis of the number of loci in the duck genome suggested that ducks have more loci than chicken. Like the quail, there appear to be many MHC class I loci, though their contributions to immune function are unknown. Our analysis of

the expressed cDNA allows us to conclude that, despite the presence of additional diversity in the genome, ducks predominantly express only one locus.

Amongst the expressed cDNA sequences, we could not identify by percentage similarity if there were alleles from a single locus. Even the highly conserved exon 4 sequences did not show a higher percentage identity between alleles U*02 and U*03 which were later found to be from the same locus. Inter-locus gene conversion is one mechanism of generating diversity at functional MHC class I loci. It is possible that inter-locus gene conversion in the duck could be responsible for the similarites between alleles from different loci. The sequence gU*12, with identity across different regions to the alleles U*02 and U*03 (Figure A 1), presents a good case for inter-locus gene conversion.

Maintaining functional variation in the MHC class I alleles is important for the health and resistance of the population to pathogens. Recombination between alleles at a locus, and between alleles from different loci, help maintain MHC variation in a population. In many species studied, non-functional MHC class I genes, or pseudogenes, exist. These are available for recombination with the functional loci to generate new alleles. The additional loci in the duck may contribute to the diversity of the expressed locus through recombination.

The Galloanseriform lineage, including waterfowl (ducks and geese) and fowl (chicken and quail) diverged from the ancestral avian lineage100 million years ago (mya). Ducks and fowl diverged from each other 90 mya (van Tuinen and Hedges, 2001), while chicken and quail diverged from each other 40 mya. Duck MHC appears to be similar to the chicken, having dominant expression of only one MHC class I locus,

which is linked to the TAP2. However, it differs from chicken, and is more like quail, in having more class I loci. Complete sequencing of the duck MHC will reveal the organization of this regionin ducks.

5 REFERENCES

Afanassieff, M., Goto, R. M., Ha, J., Sherman, M. A., Zhong, L., Auffray, C., Coudert, F., Zoorob, R., and Miller, M. M. 2001. At least one class I gene in restriction fragment pattern-Y (Rfp-Y), the second MHC gene cluster in the chicken, is transcribed, polymorphic, and shows divergent specialization in antigen binding region. *Journal of Immunology.* 166: 3324-33.

Ahluwalia, N., Bergeron, J. J., Wada, I., Degen, E., and Williams, D. B. 1992. The p88 molecular chaperone is identical to the endoplasmic reticulum membrane protein, calnexin. *Journal of Biological Chemistry.* 267: 10914-8.

Akira, S., Okazaki, K., and Sakano, H. 1987. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science.* 238: 1134-1138.

Alt, F. and Baltimore, D. 1982. Joining of immunoglobulin heavy chain gene segments: Implication from a chromosome with evidence of three D-J_H fusions. *Proceedings of the National Academy of Sciences of the United States of America.* 79: 4118-4122.

Antao, A. B., Chinchar, V. G., McConnell, T. J., Miller, N. W., Clem, L. W., and Wilson, M. R. 1999. MHC class I genes of the channel catish: sequence analysis and expression. *Immunogenetics.* 49: 303-311.

Antica, M., Wu, L., Shortman, K., and Scollay, R. 1994. Thymic stem cells in mouse bone marrow. *Blood.* 84: 111-117.

Aoyagi, K., Dijkstra, J. M., Xia, C., Denda, I., Ototake, M., Hashimoto, K., and Nakanishi, T. 2002. Classical MHC class I genes composed of highly divergent sequence lineages share a single locus in rainbow trout (*Oncorhynchus mykiss*). *The Journal of Immunology.* 168: 260-273.

Arden, B. and Klein, J. 1982. Biochemical comparison of major histocompatibility complex molecules from different subspecies of *Mus musculus*: Evidence for *trans*-specific evolution of alleles. *Proceedings of the National Academy of Sciences of the United States of America.* 79: 2342-2346.

Auchincloss, H., Sykes, M., and Sachs, D. H. 1999. Transplantation Immunology. *In* W. E. Paul (ed.): *Fundamental Immunology*, Lippincott Raven, Philadelphia.

Augustin, A., Kubo, R. T., and Sim, G. 1989. Resident pulmonary lymphocytes expressing the gd T-cell receptor. *Nature.* 340: 239-241.

Aultman, D., Adamashvili, I., Yaturu, K., Langford, M., Gelder, F., Gautreaux, M., Ghali, G. E., and McDonald, J. 1999. Soluble HLA in human body fluids. *Human Immunology.* 60: 239-244.

Ayala, F. J. 1995. The myth of Eve: molecular biology and human origins. *Science.* 270: 1930-1936.

Barber, I., Arnott, S. A., Braithwaite, V. A., Andrew, J., and Huntingford, F. A. 2001. Indirect finess consequences of mate choice in sticklebacks: offspring of brighter males grow slowly but resist parasitic infections. *Proceedings of the Royal Society of London- Series B: Biological Sciences.* 268: 71-76.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. *Proceedings of the National Academy of Sciences of the United States of America.* 91: 2216-20.

Bartl, S., Baish, M. A., Flajnik, M. F., and Ohta, Y. 1997. Identification of class I genes in cartilaginous fish, the most ancient group of vertebrates displaying an adaptive immune response. *The Journal of Immunology.* 159: 6097-104.

Bauer, S., Groh, V., Wu, J., Steinle, A., Phillips, J. H., Lanier, L. L., and Spies, T. 1999. Activation of NK cells and T cells by NKG2D, a receptor for stress-inducible MICA. *Science.* 285: 727-729.

Beauchamp, G. K. and Yamazaki, K. 1997. HLA and mate selection in humans: commentary. *American Journal of Human Genetics.* 61: 494-496.

Benoist, C. and Mathis, D. 1999. T-Lymphocyte Differentiation and Biology. *In* W. E. Paul (ed.): *Fundamental Immunology*, Lippincott Raven, Philadelphia.

Bernot, A., Zoorob, R., and Auffray, C. 1994. Linkage of a new member of the lectin supergene family to chicken *Mhc* genes. *Immunogenetics.* 39: 221-229.

Betz, U. A. K., Mayer, W. E., and Klein, J. 1994. Major histocompatibility complex class I genes of the coelacanth *Latimeria chalunmae*. *Proceedings of the National Academy of Sciences of the United States of America.* 91: 11065-11069.

Bevan, M. J. 1977. In a radiation chimera, host H-2 antigens determine immune responsiveness of donor cytotoxic cells. *Nature.* 269: 417-418.

Biassoni, R., Cantoni, C., Pende, D., Sivori, S., Parolini, S., Vitale, M., Bottino, C. and Moretta, A. 2001. Human natural killer cell receptors and co-receptors. *Immunological Reviews.* 181: 203-214.

Bingulac-Popovic, J., Figueroa, F., Sato, A., Talbot, W. S., Johnson, S. L., Gates, M., Postlethwait, J. H., and Klein, J. 1997. Mapping of *Mhc* class I and class II regions to different linkage groups in the zebrafish, *Danio rerio*. *Immunogenetics.* 46: 129-134.

Bjorkman, P. J., and Parham, P. 1990. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annual Review of Biochemistry.* 59: 253-288.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C. 1987a. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature.* 329: 512-8.

Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., and Wiley, D. C. 1987b. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature.* 329: 506-512.

Bjorkman, P. J., Strominger, J. L., and Wiley, D. C. 1985. Crystallization and X-ray diffraction studies on the histocompatibility antigens HLA-A2 and HLA-A28 from human cell membranes. *Journal of Molecular Biology.* 186: 205-10.

Black, F. L. and Hedrick, P. W. 1997. Strong balancing selecton at HLA loci: evidence from segregation in South Amerindian families. *Proceedings of the National Academy of Sciences of the United States of America.* 94: 12452-12456.

Bloom, S. E. and Bacon, L. D. 1985. Linkage of the major histocompatibility (B) complex and the nucleolar organizer in the chicken. *The Journal of Heredity.* 76: 146-154.

Bodmer, W. F. 1972. Evolutionary significance of the HL-A system. *Nature.* 237: 139-145.

Borrego, F., Ulbrecht, M., Weiss, E. H., Coligan, J. E., and Brooks, A. G. 1998. Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. *The Journal of Experimental Medicine.* 187: 813-818.

Boyd, L. F., Kozlowski, S., and Margulies, D. H. 1992. Solution binding of an antigenic peptide to a major histocompatibility complex class I molecule and the role of b 2-microglobulin. *Proceedings of the National Academy of Sciences of the United States of America.* 89: 2242-2246.

Braud, V. M., Allan, D. S. J., O'Callaghan, C., A., Soderstrom, K., D'Andrea, A., Ogg, G. S., Lazetic, S., Young, N. T., Bell, J. I., Phillips, J. H., Lanier, L. L., and McMichael, A. J. 1998. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature.* 391: 795-799.

Brennan, J., Mahon, G., Mager, D. L., Jefferies, W. A., and Takei, F. 1996. Recognition of class I major histocompatibility complex molecules by Ly-49: Specificities and domain interactions. *Journal of Experimental Medicine.* 183: 1553-1559.

Briles, W. E., Stone, H. A., and Cole, R. K. 1977. Marek's Disease: effects of B histocompatibility alloalleles in resistant and susceptibile chicken lines. *Science.* 195: 193-195.

Bucy, R. P., Chen, C. H., and Cooper, M. D. 1991. Analysis of gd T cells in the chicken. *Seminars in Immunology.* 3: 109-117.

Bukowski, J. F., Morita, C. T., and Brenner, M. B. 1994. Recognition and destruction of virus-infected cells by human gd CTL. *The Journal of Immunology.* 153: 5133-5140.

Bumstead, N. 1998. Genomic mapping of resistance to Marek's disease. *Avian Pathology.* 27: S78-S81.

Carrington, M., Nelson, G. W., Martin , M. P., Kissner, T., Vlahov, D., Goedert, J. J., Kaslow, R., Buchbinder, S., Hoots, K., and O'Brien, S. J. 1999. HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage. *Science.* 283: 1748-1752.

Carroll, L. S., Penn, D. J., and Potts, W. K. 2002. Discrimination of MHC-derived odors by untrained mice is consistent with divergence in peptide-binding region residues. *Proceedings of the National Academy of Sciences of the United States of America.* 99: 2187-92.

Cerwenka, A., Bakker, A. B. H., McClanahan, T., Wagner, J., Wu, J., Phillips, J. H., and Lanier, L. L. 2000. *Retinoic acid early* inducible genes define a ligand family for the activating NKG2D receptor in mice. *Immunity.* 12: 721-727.

Chien, Y., Iwashima, M., Kaplan, K. B., Elliott, J. F., and Davis, M. M. 1987a. A new T-cell receptor gene located within the alpha locus and expressed early in T-cell differentiation. *Nature.* 327: 677-682.

Chien, Y., Iwashima, M., Wettstein, D. A., Kaplan, K. B., Elliott, J. F., Born, W., and Davis, M. M. 1987b. T-cell receptor g gene rearrangements in early thymocytes. *Nature.* 330: 722-727.

Clark, M. S., Shaw, L., Kelly, A., Snell, P., and Elgar, G. 2001. Characterization of the MHC class I region of the Japanese pufferfish (*Fugu rubripes). Immunogenetics.* 52: 174-185.

Curtis, A. S. G. and Rooney, P. 1979. H-2 restriction of contact inhibition of epithelial cells. *Nature.* 281: 222-223.

Davis, M. M. and Bjorkman, P. J. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature.* 334: 395-402.

Davis, M. M., and Chien, Y. 1999. T-Cell Antigen Receptors. *In* W. E. Paul (ed.): *Fundamental Immunology*, Lippincott Raven, Philadelphia.

Deverson, E. V., Leong, L., Seelig, A., Coadwell, W. J., Tredgett, E. M., Butcher, G. W., and Howard, J. C. 1998. Functional analysis by site-directed mutagenesis of the complex polymorphism in rat transporter associated with antigen processing. *Journal of Immunology.* 160: 2767-79.

Doherty, P. C. and Zinkernagel, R. M. 1975. Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature.* 256: 50-52.

Due, C., Simonsen, M., and Olsson, L. 1986. The major histocompatibility complex class I heavy chain as a structural subunit of the human cell membrane insulin receptor: Implications for the range of biological functions of histocompatibility antigens. *Proceedings of the National Academy of Sciences of the United States of America.* 83: 6007-6011.

Edwards, S. V., Gasper, J., Garrigan, D., Martindale, D., and Koop, B. F. 2000. A 39-kb sequence around a blackbird *Mhc* class II gene: Ghost of selection past and songbird genome architecture. *Molecular Biology and Evolution.* 17: 1384-1395.

Edwards, S. V., Gasper, J., and March, M. 1998. Genomics and polymorphism of Agph-DAB1, an *Mhc* class II B gene in red-winged blackbirds (*Agelaius phoeniceus*). *Molecular Biology and Evolution.* 15: 236-250.

Edwards, S. V., Hess, C. M., Gasper, J., and Garrigan, D. 1999. Toward an evolutionary genomics of the avian Mhc. *Immunological Reviews.* 167: 119-132.

Ehlers, A., Beck, S., Forbes, S. A., Trowsdale, J., Volz, R., R., Y., and Ziegler, A. 2000. MHC-Linked Olfactory Receptor Loci Exhibit Polymorphism and Contribute to Extended HLA/OR-Haplotypes. *Genome Research.* 10: 1968-1978.

Ekblom, R., Grahn, M., and Hoglund, J. 2003. Patterns of polymorphism in the MHC class II of a non-passerine bird, the great snipe (*Gallinago media*). *Immunogenetics.* 54: 734-741.

Engelhard, V. H. 1994. Structure of peptides associated with MHC class I molecules. *Current Opinion in Immunology.* 6: 13-23.

Falk, K., Rotzschke, O., Stevanovi, S., Jung, G., and Rammensee, H. G. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature.* 351: 290-6.

Feinberg, A. P. and Vogelstein, B. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry.* 132: 6-13.

Flajnik, M. F., Kasahara, M., Shum, B. P., Salter-Cid, L., Taylor, E., and Du Pasquier, L. 1993. A novel type of class I gene organization in vertebrates: a large family of non-MHC-linked class I genes is expressed at the RNA level in the amphibian *Xenopus*. *EMBO Journal.* 12: 4385-96.

Flajnik, M. F., Ohta, Y., Greenberg, A. S., Salter-Cid, L., Carrizosa, A., Du Pasquier, L., and Kasahara, M. 1999a. Two ancient allelic lineages at the single classical class I locus in the Xenopus MHC. *The Journal of Immunology.* 163: 3826-33.

Flajnik, M. F., Ohta, Y., Namikawa-Yamada, C., and Nonaka, M. 1999b. Insight into the primordial MHC from studies in ectothermic vertebrates. *Immunological Reviews.* 167: 59-67.

Flajnik, M. F. and Kasahara, M. 2001 Comparative Genomics of the MHC: Glimpses into the Evolution of the adaptive immune system.

Freeman-Gallant, C. R., Johnson, E. M., Saponara, F., and Stanger, M. 2002. Variation at the major histocompatibility complex in Savannah sparrows. *Molecular Ecology.* 11: 1125-1130.

Gaczynska, M., Rock, K. L., Speis, T., and Goldberg, A. L. 1994. Peptidase activities of proteasomes are differentially regulated by the major histocompatibility complex-encoded genes for LMP2 and LMP7. *Proceedings of the National Academy of Sciences of the United States of America.* 91: 9213-9217.

Gallon, L., Watschinger, B., Murphy, B., Akalin, E., Sayegh, M. H., and Carpenter, C. B. 1995. The indirect pathway of allorecognition. *Transplantation.* 59: 612-616.

Gao, B., Adhikari, R., Howarth, M., Nakamura, K., Gold, M. C., Hill, A. B., Knee, R., Michalak, M., and Elliott, T. 2002. Assembly and antigen-presenting function of MHC class I molecules in cells lacking the ER chaperone calreticulin. *Immunity.* 16: 99-109.

Garboczi, D. N., Ghosh, P., Utz, U., Fan, Q. R., Biddison, W. E., and Wiley, D. C. 1996. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature.* 384: 134-41.

Garrett, T. P., Saper, M. A., Bjorkman, P. J., Strominger, J. L., and Wiley, D. C. 1989. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature.* 342: 692-696.

Gascoigne, N. R. J., Chien, Y., Becker, D. M., Kavaler, J., and Davis, M. M. 1984. Genomic organization and sequence of T-cell receptor b-chain constant- and joining-region genes. *Nature.* 310: 387-391.

Gasper, J. S., Shiina, T., Inoko, H., and Edwards, S. V. 2001. Songbird genomics: analysis of 45 kb upstream of a polymorphic *Mhc* class II gene in red-winged blackbirds (*Agelaius phoeniceus*). *Genomics.* 75: 26-34.

Germain, R. N., and Margulies, D. H. 1993. The biochemistry and cell biology of antigen processing and presentation. *Annual Review of Immunology.* 11: 403-450.

Gilbert, S. C., Plebanski, M., Gupta, S., Morris, J., Cox, M., Aidoo, M., Kwiatkowski, D., Greenwood, B. M., Whittle, H. C., and Hill, A. V. S. 1998. Association of malaria parasite population structure, HLA, and immunological antagonism. *Science.* 279: 1173-1177.

Gilder, P. M. and Slater, P. J. B. 1978. Interest of mice in conspecific male odours is influenced by degree of kinship. *Nature.* 274: 364-365.

Goodman, T. and Lefrancois, L. 1988. Expression of the g-d T-cell receptor on intestinal CD8+ intraepithelial lymphocytes. *Nature.* 333: 855-858.

Gorman, O. T., Bean, W. J., Kawaoka, Y., Donatelli, I., Guo, Y. J., and Webster, R. G. 1991. Evolution of influenza A virus nucleoprotein genes: implications for the origins of H1N1 human and classical swine viruses. *Journal of Virology.* 65: 3704-14.

Grandea, A. G., Golovina, T. N., Hamilton, S. E., Sriram, V., Spies, T., Brutkiewicz, R. R., Harty, J. T., Eisenlohr, L. C., and Van Kaer, L. 2000. Impaired assembly yet normal trafficking of MHC class I molecules in Tapasin mutant mice. *Immunity.* 13: 213-22.

Grandea, A. G., 3rd, Lehner, P. J., Cresswell, P., and Spies, T. 1997. Regulation of MHC class I heterodimer stability and interaction with TAP by tapasin. *Immunogenetics.* 46: 477-83.

Grant, E. P., Michalek, M. T., Goldberg, A. L., and Rock, D. L. 1995. Rate of antigen degradation by the ubiquitin-proteasome pathway influences MHC class I presentation. *The journal of Immunology.* 155: 3750-3758.

Grimholt, U. 1997. Transport-associated proteins in Atlantic slmon (*Salmo salar*). *Immunogenetics.* 46: 213-221.

Grimholt, U., Drablos, F., Jorgensen, S. M., Hoyheim, B., and Stet, R. J. M. 2002. The major histocompatibility class I locus in Atlantic salmon (*Salmo salar* L.): polymorphism, linkage analysis and protein modelling. *Immunogenetics.* 54: 570-581.

Grimholt, U., Hordvik, I., Fosse, V. M., Olsaker, I., Endresen, C., and Lie, O. 1993. Molecular cloning of major histocompatibility complex class I cDNAs from Atlantic salmon (*Salmo salar*). *Immunogenetics.* 37: 469-473.

Groettrup, M., Ungewiss, K., Azogui, O., Palacios, R., Owen, M. J., Hayday, A. C., and von Boehmer, H. 1993. A novel disulfide-linked heterodimer on pre-T cells consists of the T cell receptor b chain and a 33 kd glycoprotein. *Cell.* 75: 283-294.

Gromme, M., and Neefjes, J. 2002. Antigen degradation or presentation by MHC class I molecules via classical and non-classical pathways. Molecular Immunology. 39: 181-202.

Guild, B. C., Erikson, R. L., and Strominger, J. L. 1983. HLA-A2 and HLA-B7 antigens are phosphorylated in vitro by rous sarcoma virus kinase (pp60v-src) at a tyrosine residue encoded in a highly conserved exon of the intracellular domain. *Proceedings of the National Academy of Sciences of the United States of America.* 80: 2894-8.

Guillemot, F., Billault, A., Pourquie, O., Behar, G., Chausse, A., Zoorob, R., Kreibich, G., and Auffray, C. 1988. A molecular map of the chicken major histocompatibility complex: the class II b genes are closelly linked to the class I genes and the nucleolar organizer. *The EMBO Journal.* 7: 2775-2785.

Gunther, E. and Walter, L. 2001. The major histocompatibility complex of the rat (*Rattus norvegicus*). *Immunogenetics.* 53: 520-542.

Hamilton, W. D. and Zuk, M. 1982. Heritable true fitness and bright birds: a role for parasites? *Science.* 218: 384-387.

Hansen, J. D., Strassburger, P., Thorgaard, G. H., Young, W. P., and Du Pasquier, L. 1999. Expression, linkage, and polymorphism of MHC-related genes in Rainbow Trout, *Oncorhynchus mykiss*. *The Journal of Immunology.* 163: 774-786.

Hayday, A. C., Diamond, D. J., Tanigawa, G., Heilig, J. S., Folsom, V., Saito, H., and Tonegasa, S. 1985. Unusual organization and diversity of T-cell receptor a-chain genes. *Nature.* 316: 828-832.

Hedrick, P. W. and Black, F. L. 1997. HLA and mate selection: no evidence in South Amerindians. *American Journal of Human Genetics.* 61: 505-511.

Henderson, R. A., Michel, H., Sakaguchi, K., Shabanowitz, J., Appella, E., Hunt, D. F., and Engelhard, V. H. 1992. HLA-A2.1-Associated Peptides from a mutant cell line: A second pathway of antigen presentation. *Science.* 255: 1264-1266.

Hess, C. M., Gasper, J., Hoekstra, H. E., Hill, C. E., and Edwards, S. V. 2000. MHC class II pseudogene and genomic signature of a 32-kb cosmid in the house finch (*Carpodacus mexicanus*). *Genome Research.* 10: 613-623.

Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., Bennett, S., Brewster, D., McMichael, A. J., and Greenwood, B. M. 1991. Common West African HLA antigens are associated with protection from sever malaria. *Nature.* 352: 595-600.

Ho, H. N., Yang, Y. S., Hsieh, R. P., Lin, H. R., Chen, S. U., Chen, H. F., Huang, S. C., Lee, T. Y., and Gill, T. G. 1994. Sharing of human leukocyte antigens in couples with unexplained infertility affects the success of in vitro fertilizaion and tubal embryo transfer. *American Journal of Obstetrics and Gynecology.* 170: 63-71.

Hochstenbach, F., David, V., Watkins, S., and Brenner, M. B. 1992. Endoplasmic reticulum resident protein of 90 kilodaltons associates with the T- and B-cell antigen receptors and major histocompatibility complex antigens during their assembly. *Proceedings of the National Academy of Sciences of the United States of America.* 89: 4734-8.

Hofmann, K. and Stoffel, W. 1993. TMbase - A database of membrane spanning proteins segments. *Biological Chemistry Hoppe-Seyler.* 374: 166.

Hughes, A. L. and Nei, M. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature.* 335: 167-170.

Hunt, H. D. and Fulton, J. E. 1998. Analysis of polymorphisms in the major expressed class I locus (B-FIV) of the chicken. *Immunogenetics.* 47: 456-67.

Hyde, S. C., Emsley, P., Hartshorn, M. J., Mimmack, M. M., Gileadi, U., Pearce, S. R., Gallagher, M. P., Gill, D. R., Hubbard, R. E., and Higgins, C. F. 1990. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport.[comment]. *Nature.* 346: 362-5.

Ignatowicz, L., Kappler, J., and Marrack, P. 1996. The repertoire of T cells shaped by a single MHC/peptide ligand. *Cell.* 84: 521-529.

Isles, A. R., Baum, M. J., Ma, D., Keverne, E. B., and Allen, N. D. 2001. Urinary odour preferences in mice. *Nature.* 409: 783-784.

Itohara, S., Farr, A. G., Lafaille, J. J., Bonneville, M., Takagaki, Y., Haas, W., and Tonegawa, S. 1990. Homing of a gd thymocyte subset with homogeneous T-cell receptors to mucosal epithelia. *Nature.* 343: 754-7557.

Jacob, S., McClintock, M. K., Zelano, B., and Ober, C. 2002. Paternally inherited HLA alleles are associated with women's choice of male odor. *Nature Genetics.* 30: 175-179.

Janeway, C., Rudensky, A., Rath, S., and Murphy, D. 1992. It is easier for a camel to pass the needle's eye. *Current Biology.* 2: 26-28.

Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R., and Wiley, D. C. 1991. Identification of self peptides bound to purified HLA-B27. *Nature.* 353: 326-9.

Jarvi, S. I., Goto, R. M., Gee, G. F., Briles, W. E., and Miller, M. M. 1999. Identification, inheritance, and linkage of B-G-like and MHC class I genes in cranes. *The Journal of Heredity.* 90: 152-159.

Joly, E. and Butcher, G. W. 1998. Why are there two rat TAPs? *Immunology Today.* 19: 580-5.

Joly, E., Deverson, E. V., Coadwell, J. W., Gunther, E., Howard, J. C., and Butcher, G. W. 1994. The distribution of Tap2 alleles among laboratory rat RT1 haplotypes. *Immunogenetics.* 40: 45-53.

Joly, E., Leong, L., Coadwell, W. J., Clarkson, C., and Butcher, G. W. 1996. The rat MHC haplotype RT1c expresses two classical class I molecules. *The Journal of Immunology.* 157: 1551-8.

Juul-Madsen, H. R., Dalgaard, T. S., Guldbrandtsen, B., and Salomonsen, J. 2000. A polymorphic major histocompatibility complex class II-like locus maps outside of both the chicken B-system and *Rfp-Y*-system. *European Journal of Immunogenetics.* 27: 63-71.

Karttunen, J. T., Lehner, P. J., Gupta, S. S., Hewitt, E. W., and Cresswell, P. 2001. Distinct functions and cooperative interaction of the subunits of the transporter associated with antigen processing (TAP). *Proceedings of the National Academy of Sciences of the United States of America.* 98: 7431-7436.

Kaufman, J. 1999. Co-evolving genes in MHC haplotypes: the "rule" for nonmammalian vertebrates? *Immunogenetics.* 50: 228-36.

Kaufman, J., Jacob, J., Shaw, I., Walker, B., Milne, S., Beck, S., and Salomonsen, J. 1999a. Gene organisation determines evolution of function in the chicken MHC. *Immunological Reviews.* 167: 101-17.

Kaufman, J., Milne, S., Gobel, T. W., Walker, B. A., Jacob, J. P., Auffray, C., Zoorob, R., and Beck, S. 1999b. The chicken B locus is a minimal essential major histocompatibility complex. *Nature.* 401: 923-5.

Kaufman, J., Salomonsen, J., and Flajnik, M. 1994. Evolutionary conservation of MHC class I and class II molecules-- different yet the same. *Seminars in Immunology.* 6: 411-424.

Kaufman, J. and Venugopal, K. 1998. The importance of MHC for Rous sarcoma virus and Marek's disease virus--some Payne-ful considerations. *Avian Pathology.* 27: S82-S87.

Kaufman, J., Volk, H., and Wallny, H. J. 1995. A "minimal essential Mhc" and an "unrecognized Mhc": two extremes in selection for polymorphism. *Immunological Reviews.* 143: 63-88.

Kavaler, J., Davis, M. M., and Chien, Y. 1984. Localization of a T-cell receptor diversity-region element. *Nature.* 310: 421-423.

Kelly, A., Powis, S. H., Kerr, L. A., Mockridge, I., Elliott, T., Bastin, J., Uchanska-Ziegler, B., Ziegler, A., Trowsdale, J., and Townsend, A. 1992a. Assembly and function of the two ABC transporter proteins encoded in the human major histocompatibility complex. *Nature.* 355: 641-4.

Kelly, A., Radley, E., Khurshid, F., Alderton, R. P., Trowsdale, J., and Ohta, Y. 1992b. Isolation of transporter associated with antigen processing genes, TAP1 and TAP2, from the horned shark *Heterodontus francisci. Journal of Molecular Biology.* 228: 433-41.

Kida, H., Yanagawa, R., and Matsuoka, Y. 1980. Duck influenza lacking evidence of disease signs and immune response. *Infection and Immunity.* 30: 547-53.

Kisielow, P., Bluthmann, H., Staerz, U. D., Steinmetz, M., and von Boehmer, H. 1988. Tolerance in T-cell-receptor transgenic mice involves deletion of nonmature $CD4^{+}8^{+}$ thymocytes. *Nature.* 333: 742-746.

Kittur, D., Shimizu, Y., DeMars, R., and Edidin, M. 1987. Insulin binding to human B lymphoblasts is a function of HLA haplotype. *Proceedings of the National Academy of Sciences of the United States of America.* 84: 1351-1355.

Kleijmeer, M. J., Kelly, A., Geuze, H. J., Slot, J. W., Townsend, A., and Trowsdale, J. 1992. Location of MHC-encoded transporters in the endoplasmic reticulum and cis-Golgi. *Nature.* 357: 342-344.

Koning, F., Stingl, G., Yokoyama, W. M., Yamada, H., Maloy, W. L., Tschachler, E., Shevach, E. M., and Coligan, J. E. 1987. Identification of a T3-associated gd T cell receptor on Thy-1+ dendritic epidermal cell lines. *Science.* 236: 834-837.

Kostyu, D. D., Dawson, D. V., Elias, S., and Ober, C. 1993. Deficit of HLA homozygotes in a caucasian isolate. *Human Immunology.* 37: 135-142.

Kroemer, G., Zoorob, R., and Auffray, C. 1990. Structure and expression of a chicken MHC class I gene. *Immunogenetics.* 31: 405-9.

Kuhner, M., Watts, S., Klitz, W., Thomson, G., and Goodenow, R. S. 1990. Gene conversion in the evolution of both the H-2 and Qa class I genes of the murine major histocompatibility complex. *Genetics.* 126: 1115-1126.

Kulski, J., Shiina, T., Anzai, T., Kohara, S., and Inoko, H. 2002. Comparative genomic analysis of the MHC: the evolution of class I duplication blocks, diversity and complexity from shark to man. *Immunological Reviews.* 190: 95-122.

Kumar, S. and Hedges, S. B. 1998. A molecular timescale for vertebrate evolution. *Nature.* 392: 917-920.

Landry, C., Garant, D., Duchesne, P., and Bernatchez, L. 2001. 'Good genes as heterozygosity': the major histocompatibility complex and mate choice in Atlantic salmon (*Salmo salar*). *Proceedings of the Royal Society of London-Series B: Biological Sciences.* 268: 1279-1285.

Lanier, L. L., Borliss, B., Wu, J., and Phillips, J. H. 1998. Association of DAP12 with activating CD94/NKG2C NK cell receptors. *Immunity.* 8: 693-701.

Laver, W. G. and Webster, R. G. 1973. Studies on the origin of pandemic influenza: evidence implicating duck and equine influenza viruses as possible progenitors of the Hong Kong strain of human influenza. *Virology.* 51: 383-391.

Lawlor, D. A., Ward, F. E., Ennis, P. D., Jackson, A. P., and Parham, P. 1988. HLA-A and B polymorphisms predate the divergence of humans and chimpanzees. *Nature.* 335: 268-71.

Lawlor, D. A., Zemmour, J., Ennis, P. D., and Parham, P. 1990. Evolutionn of class I MHC genes and proteins: from natural selection to thymic selection. *Annual Review of Immunology.* 8: 23-63.

Lazetic, S., Chang, C., Houchins, J. P., Lanier, L. L., and Phillips, J. H. 1996. Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits. *The Journal of Immunology.* 157: 4741-4745.

LePage, K. T., Miller, M. M., Briles, W. E., and Taylor, R. L. 2000. Rfp-Y genotype affects the fate of Rous sarcomas in B2B5 chickens. *Immunogenetics.* 51: 751-754.

Loconto, J., Papes, F., Chang, E., Stowers, L., Jones, E. P., Takada, T., Kumanovics, A., Fischer Lindahl, K., and Dulac, C. 2003. Functional expression of murine V2R pheromone receptors involves selective association with the M10 and M1 families of MHC class 1b molecules. *Cell.* 112: 607-618.

Lohwasser, S., Kubota, A., Salcedo, M., Lian, R. H., and Takei, F. 2001. The non-classical MHC class I molecule Qa-1b inhibits classical MHC class I-restricted cytotoxicity of cytotoxic T lymphocytes. *International Immunology.* 13: 321-327.

Loo, T. W., Bartlett, M. C., and Clarke, D. M. 2002. The "LSGGQ" motif in each nucleotide-binding domain of human P-glycoprotein is adjacent to the opposing walker A sequence. *Journal of Biological Chemistry.* 277: 41303-6.

Madden, D. R., Gorga, J. C., Strominger, J. L., and Wiley, D. C. 1991. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature.* 353: 321-5.

Manning, C. J., Wakeland, E. K., and Potts, W. K. 1992. Communal nesting patterns in mice implicate MHC genes in kin recognition. *Nature.* 360: 581-583.

Mason, W. S., Seal, G., and Summers, J. 1980. Virus of pekin ducks with structural and biological relatedness to human hepatitis B virus. *Journal of Virology.* 36: 829-836.

Matsuo, M. Y., Asakawa, S., Shimizu, N., Kimura, H., and Nonaka, M. 2002. Nucleotide sequence of the MHC class I genomic region of a teleost the medaka (*Oryzia latipes*). *Immunogenetics.* 53: 930-940.

Matzinger, P., Zamoyska, R., and Waldmann, H. 1984. Self tolerance is H-2-restricted. *Nature.* 308: 738-41.

MHC Sequencing Consortium. 1999. Complete sequence and gene map of a human major histocompatibility complex. *Nature.* 401: 921-923.

Miller, M. M., Goto, R., Bernot, A., Zoorob, R., Auffray, C., Bumstead, N., and Briles, W. E. 1994. Two MHC class I and two MHC class II genes map to the chicken *Rfp-Y* system outside the B complex. *Proceedings of the National Academy of Sciences of the United States of America.* 91: 4397-401.

Miller, M. M., Goto, R. M., Taylor, R. L., Zoorob, R., Auffray, C., Briles, R. W., Briles, W. E., and Bloom, S. E. 1996. Assignment of *Rfp-Y* to the chicken major histocompatibility complex/NOR microchromosome and evidence for high-frequency recombination associated with the nucleolar organizer region. *Proceedings of the National Academy of Sciences of the United States of America.* 93: 3958-3962.

Momburg, F., Armandola, E. A., Post, M., and Hammerling, G. J. 1996. Residues in TAP2 peptide transporters controlling substrate specificity. *The Journal of Immunology.* 156: 1756-63.

Momburg, F., Roelse, J., Howard, J. C., Butcher, G. W., Hammerling, G. J., and Neefjes, J. J. 1994. Selectivity of MHC-encoded peptide transporters from human, mouse and rat. *Nature.* 367: 648-51.

Montag, S., Frank, M., Ulmer, H., Wernet, D., Gopel, W., and Rammensee, H. G. 2001. "Electronic nose" detects major histocompatibility complex-dependent prerenal and postrenal odor components. *Proceedings of the National Academy of Sciences of the United States of America.* 98: 9249-54.

Namikawa, C., Salter-Cid, L., Flajnik, M. F., Kato, Y., Nonaka, M., and Sasaki, M. 1995. Isolation of Xenopus LMP-7 homologues. *The Journal of Immunology.* 155: 1964-1971.

Naruse, K., Fukamachi, S., Mitani, H., Kondo, M., Matsuoka, T., Kondo, S., Hanamura, N., Morita, Y., Hasegawa, K., Nishigaki, R., Shimada, A., Wada, H., Kusakabe, T., Suzuki, N., Kinoshita, M., Kanamor, A., Terado, T., Kimura, H., Nonaka, M., and Shima, A. 2000. A detailed linkage map of Medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics.* 154: 1773-1784.

Neefjes, J. J., Momburg, F., and Hammerling, G. J. 1993. Selective and ATP-dependent translocation of peptides by the MHC-encoded transporter. *Science.* 261: 769-71.

Nelson, J. S. 1994. Fishes of the World. John Wiley and Sons, Inc, Toronto.

Nijenhuis, M. and Hammerling, G. J. 1996. Multiple regions of the transporter associated with antigen processing (TAP) contribute to its peptide binding site. *The Journal of Immunology.* 157: 5467-77.

Nonaka, M., Namikawa, C., Kato, Y., Sasaki, M., Salter-Cid, L., and Flajnik, M. F. 1997. Major histocompatibility complex gene mapping in the amphibian Xenopus implies a primordial organization. *Proceedings of the National Academy of Sciences of the United States of America.* 94: 5789-91.

Nonaka, M., Yamada-Namikawa, C., Flajnik, M. F., and Du Pasquier, L. 2000. Trans-species polymorphism of the major histocompatibility complex-encoded proteasome subunit LMP7 in an amphibian genus, *Xenopus*. *Immunogenetics.* 51: 186-192.

Norment, A. M., Salter, R. D., Parham, P., Engelhard, V. H., and Littman, D. R. 1988. Cell-cell adhesion mediated by CD8 and MHC class I molecules. *Nature.* 336: 79-81.

Novotny, J., Tonegawa, S., Saito, H., Kranz, D. M., and Eisen, H. N. 1986. Secondary, tertiary, and quaternary structure of T-cell -specific immunoglobulin-like polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America.* 83: 742-746.

Ober, C., Weitkamp, L. R., Cox, N., Cytch, H., Kostyu, D., and Elias, S. 1997. HLA and mate choice in humans. *American journal of human genetics.* 61: 497-504.

Ochoa, G. and Jaffe, K. 1999. On sex, mate selection and the Red Queen. *Journal of Theoretical Biology.* 199: 1-9.

Ohta, Y., McKinney, E. C., Criscitiello, M. F., and Flajnik, M. F. 2002. Proteasome, transporter associated with antigen processing, and class I genes in the nurse shark *Ginglymostoma cirratum*: evidence for a stable class I region and MHC haplotype lineages. *The Journal of Immunology.* 168: 771-81.

Ohta, Y., Okamura, K., McKinney, E. C., Bartl, S., Hashimoto, K., and Flajnik, M. F. 2000. Primitive synteny of vertebrate major histocompatibility complex class I and class II genes. *Proceedings of the National Academy of Sciences of the United States of America.* 97: 4712-7.

Ohta, Y., Powis, S. J., Coadwell, W. J., Haliniewski, D. E., Liu, Y., Li, H., and Flajnik, M. F. 1999. Identification and genetic mapping of *Xenopus* TAP2 genes. *Immunogenetics.* 49: 171-82.

Okamura, K., Ototake, M., Nakanishi, T., Kurosawa, Y., and Hashimoto, K. 1997. The most primitive vertebrates with jaws possess highly polymorphic MHC class I genes comparable to those of humans. *Immunity.* 7: 777-790.

Ortmann, B., Androlewicz, M. J., and Creswell, P. 1994. MHC class I/b2-microglobulin complexes associate with TAP transporters before peptide binding. *Nature.* 368: 864-867.

Ortmann, B., Copeman, J., Lehner, P. J., Sadasivan, B., Herberg, J. A., Grandea, A. G., Riddell, S. R., Tampe, R., Spies, T., Trowsdale, J., and Cresswell, P. 1997. A critical role for tapasin in the assembly and function of multimeric MHC class I-TAP complexes. *Science.* 277: 1306-1309.

Pamer, E., and Cresswell, P. 1998. Mechanisms of MHC class I-restricted antigen processing. *Annual Review of Immunology.* 16:323-358.

Pamer, E. G., Wang, C. R., Flaherty, L., Fischer Lindahl, K., and Bevan, M. J. 1992. H-2M3 presents a *Listeria monocytogenes* peptide to cytotoxic T lymphocytes. *Cell.* 70: 215-223.

Pardoll, D. M., Fowlkes, B. J., Bluestone, J. A., Kruisbeek, A., Maloy, W. L., Coligan, J. E., and Schwartz, R. H. 1987. Differential expression of two distinct T-cell receptors during thymocyte development. *Nature.* 326: 79-81.

Parham, P., Lawlor, D. A., Lomen, C. E., and Ennis, P. D. 1989a. Diversity and diversification of HLA-A, B, C alleles: A panel of unique HLA-A2 mutant molecules define epitopes recognized by HLA-A2-specific antibodies and cytotoxic T lymphocytes. *The Journal of Immunology.* 142: 3937-3950.

Parham, P., Lomen, C. E., Lawlor, D. A., Ways, J. P., Holmes, N., Coppin, H. L., Salter, R. D., Wan, A. M., and Ennis, P. D. 1988. Nature of polymorphism in HLA-A, -B, and -C molecules. *Proceedings of the National Academy of Sciences of the United States of America.* 85: 4005-9.

Parham, P. and Ohta, T. 1996. Population biology of antigen presentation by MHC class I molecules. *Science.* 272: 67-74.

Pazderka, F., Longenecker, B. M., Law, G. R. J., Stone, H. A., and Ruth, R. F. 1975. Histocompatibility of chicken populations selected for resistance to Marek's Disease. *Immunogenetics.* 2: 93-100.

Penn, D. and Potts, W. 1998. MHC-disassortative mating preferences reversed by cross-fostering. *Proceedings of the Royal Society of London - Series B: Biological Sciences.* 265: 1299-1306.

Persson, A., Stet, R. J. M., and Pilstrom, L. 1999. Characterization of MHC class I and b_2-microglobulin sequences in Atlantic cod reveals an unusually high number of expressed class I genes. *Immunogenetics.* 50: 49-59.

Persson, B. and Argos, P. 1994. Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *Journal of Molecular Biology.* 237: 182-92.

Persson, B. and Argos, P. 1996. Topology prediction of membrane proteins. *Protein Science.* 5: 363-71.

Peterson, P. A., Cunningham, B. A., Berggard, I., and Edelman, G. M. 1972. b2-microglobulin-A Free Immunoglobulin Domain. *Proceedings of the National Academy of Sciences of the United States of America.* 69: 1697-1701.

Peterson, P. A., Rask, L., and Lindblom, J. B. 1974. Highly purified papain-solubilized HL-A antigens contain b2-microglobulin. *Proceedings of the National Academy of Sciences of the United States of America.* 71: 35-39.

Plachy, J. 1984. Hierarchy of the B (MHC) haplotypes controlling resistance to Rous Sarcomas in a model of inbred lines of chickens. *Folia Biologica (Praha).* 30: 412-425.

Plachy, J., Pink, J. R. L., and Hala, K. 1992. Biology of the chicken MHC (B complex). *Critical Reviews in Immunology.* 121: 47-79.

Pober, J. S., Guild, B. C., and Strominger, J. L. 1978. Phosphorylation in vivo and in vitro of human histocompatibility antigens (HLA-A and HLA-B) in the carboxy-terminal intracellular domain. *Proceedings of the National Academy of Sciences of the United States of America.* 75: 6002-6.

Potts, W. K., Manning, C. J., and Wakeland, E. K. 1991. Mating patterns in seminatural populations of mice influenced by MHC genotype.[comment]. *Nature.* 352: 619-21.

Potts, W. K. and Slev, P. R. 1995. Pathogen-based models favoring MHC genetic diversity. *Immunological Reviews.* 143: 181-97.

Powis, S. J., Deverson, E. V., Coadwell, W. J., Ciruela, A., Huskisson, N. S., Smith, H., Butcher, G. W., and Howard, J. C. 1992. Effect of polymorphism of an MHC-linked transporter on the peptides assembled in a class I molecule. *Nature.* 357: 211-215.

Powis, S. J., Howard, J. C., and Butcher, G. W. 1991. The major histocompatibility complex class II-linked *cim* locus controls the kinetics of intracellular transport of a classical class I molecule. *Journal of Experimental Medicine.* 173: 913-921.

Powis, S. J., Young, L. L., Joly, E., Barker, P. J., Richardson, L., Brandt, R. P., Melief, C. J., Howard, J. C., and Butcher, G. W. 1996. The rat cim effect: TAP allele-dependent changes in a class I MHC anchor motif and evidence against C-terminal trimming of peptides in the ER. *Immunity.* 4: 159-65.

Prosite. 2002. ABC transporters family signature. , Swiss Institute of Bioinformatics.

Rajagopalan, S. and Brenner, M. B. 1994. Calnexin retains unassembled major histocompatibility complex class I free heavy chains in the endoplasmic reticulum. *Journal of Experimental Medicine.* 180: 407-412.

Reusch, T. B. H., Haberli, M. A., Aeschlimann, P. B., and Milinski, M. 2001. Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. *Nature.* 414: 300-302.

Robinson, J., Malik, A., Parham, P., Bodmer, J. G., and Marsh, S. G. E. 2000. IMGT/HLA Database-a sequence database for the human major histocompatibility complex. *Tissue Antigens.* 55: 280-287.

Rocha, B. and von Boehmer, H. 1991. Peripheral selection of the T-cell repertoire. *Science.* 251: 1225-1228.

Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwand, D., and Goldberg, A. L. 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell.* 78: 761-771.

Rogers, S., Shaw, I., Ross, N., Nair, V., Rothwell, L., Kaufman, J., and Kaiser, P. 2003. Analysis of part of the chicken Rfp-Y region reveals two novel lectin genes, the first complete genomic sequence of a class I a-chain gene, a truncated class II b-chain gene, and a large CR1 repeat. *Immunogenetics.* .

Sadasivan, B., Lehner, P. J., Ortmann, B., Spies, T., and Cresswell, P. 1996. Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity.* 5: 103-114.

Sakenell, P. S., Miller, M. M., Goto, R. M., Gauderman, W. J., and Briles, W. E. 1996. Association between the Rfp-Y haplotype and the incidence of Marek's disease in chickens. *Immunogenetics.* 44: 242-245.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.

Sammut, B., Du Pasquier, L., Ducoroy, P., Laurens, V., Marcuz, A., and Tournefier, A. 1999. Axolotl MHC architecture and polymorphism. *European Journal of Immunology.* 29: 2897-2907.

Sammut, B., Marcuz, A., and Pasquier, L. D. 2002. The fate of duplicated major histocompatibility complex class Ia genes in a dodecaploid amphibian, *Xenopus ruwenzoriensis*. *European Journal of Immunology.* 32: 1593-604.

Saper, M. A., Bjorkman, P. J., and Wiley, D. C. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *Journal of Molecular Biology.* 219: 277-319.

Sato, A., Figueroa, F., O'hUignin, C., Reznick, D. N., and Klein, J. 1995. Identification of major histocompatibility complex genes in the guppy, *Poecilia reticulata*. *Immunogenetics.* 43: 38-49.

Sato, A., Klein, D., Sultmann, H., Figueroa, F., O'hUigin, C., and Klein, J. 1997. Class I *Mhc* genes of cichlid fishes: identification, expression, and polymorphism. *Immunogenetics.* 46: 63-72.

Saveanu, L., Daniel, S., and van Endert, P. M. 2001. Distinct functions of the ATP binding cassettes of transporters associated with antigen processing. *The Journal of Biological Chemistry.* 276: 22107-22113.

Schierman, L. W., Watanabe, D. H., and McBride, R. A. 1977. Genetic control of Rous Sarcoma regression in chickens: linkage with the Major Histocompatibility Complex. *Immunogenetics.* 5: 325-332.

Schild, H., Mavaddat, N., Litzenberger, C., Ehrich, E. W., Davis, M. M., Bluestone, J. A., Matis, L., Draper, R. K., and Y., C. 1994. The nature of major histocompatibility complex recognition by gd T cells. *Cell.* 76: 29-37.

Schneider, E. and Hunke, S. 1998. ATP-binding -cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiology Reviews.* 22: 1-20.

Shawar, S. M., Vyas, J. M., Rodgers, J. R., and Rich, R. R. 1994. Antigen presentation by major histocompatibility complex class I-b molecules. *Annual Review of Immunology.* 12: 839-880.

Shiina, T., Ando, A., Imanishi, T., Kawata, H., Hanzawa, K., Gojobori, T., Inoko, H., and Watanabe, S. 1995. Isolation and characterization of cDNA clones for Japanese quail (*Coturnix japonica*) major histocompatibility complex (MhcCoja) class I molecules. *Immunogenetics.* 42: 213-6.

Shiina, T., Oka, A., Imanishi, T., Hanzawa, K., Gojobori, T., Watanabe, S., and Inoko, H. 1999a. Multiple class I loci expressed by the quail Mhc. *Immunogenetics.* 49: 456-60.

Shiina, T., Shimizu, C., Oka, A., Teraoka, Y., Imanishi, T., Gojobori, T., Hanzawa, K., Watanabe, S., and Inoko, H. 1999b. Gene organization of the quail major histocompatibility complex (MhcCoja) class I gene region. *Immunogenetics.* 49: 384-94.

Shum, B. P., Avila, D., Du Pasquier, L., Kasahara, M., and Flajnik, M. F. 1993. Isolation of a classical MHC class I cDNA from an amphibian. Evidence for only one class I locus in the Xenopus MHC. *J Immunol.* 151: 5376-86.

Shum, B. P., Guethlein, L., Flodin, L. R., Adkison, M. A., Hedrick, R. P., Nehring, R. B., Stet, R. J., Secombes, C., and Parham, P. 2001. Modes of salmonid MHC class I and II evolution differ from the primate paradigm. *The Journal of Immunology.* 166: 3297-308.

Shum, B. P., Rajalingam, R., Magor, K. E., Azumi, K., Carr, W. H., Dixon, B., Stet, R. J., Adkison, M. A., Hedrick, R. P., and Parham, P. 1999. A divergent non-classical class I gene conserved in salmonids. *Immunogenetics.* 49: 479-90.

Simister, N. E. and Mostov, K. E. 1989. An Fc receptor structurally related to MHC class I antigens. *Nature.* 337: 184-187.

Singer, A. G., Beauchamp, G. K., and Yamazaki, K. 1997. Volatile signals of the major histocompatibility complex in male mouse urine. *Proceedings of the National Academy of Sciences of the United States of America.* 94: 2210-2214.

Singer, A. G., Tsuchiya, H., Wellington, J. L., Beauchamp, G. K., and Yamazaki, K. 1993. Chemistry of odortypes in mice: fractionation and bioassay. *Journal of Chemical Ecology.* 19: 569-579.

Singh, P. B., Brown, R. E., and Roser, B. 1987. MHC antigens in urine as olfactory recognition cues. *Nature.* 327: 161-4.

Skjodt, K., Koch, C., Crone, M., and Simonsen, M. 1985. Analysis of chickens for recombination within the MHC (B-complex). *Tissue Antigens.* 25: 278-282.

Slade, R. W. and McCallum, H. I. 1992. Overdominant vs. frequency-dependent selection at MHC loci. *Genetics.* 132: 861-862.

Sobst, B., Zavazava, N., Luszyk, D., Lange, K., Ussat, S., Eggert, F., Ferstl, R., and Muller-Ruchholtz, W. 1999. Molecular forms of soluble HLA in body fluids: potential determinants of body odor cues. *Genetica.* 104: 275-283.

Solheim, J. C., Harris, M. R., Kindle, C. S., and Hansen, T. H. 1997. Prominence of b 2-microglobulin, class I heavy chain conformation, and tapasin in the interactions of class I heavy chain with calreticulin and the transporter associated with antigen processing. *The Journal of Immunology.* 158: 2236-2241.

Spies, T., Cerundolo, V., Colonna, M., Cresswell, P., Townsend, A., and DeMars, R. 1992. Presentation of viral antigen by MHC class I molecules is dependent on a putative peptide transporter heterodimer. *Nature.* 355: 644-6.

Spits, H., Lanier, L. L., and Phillips, J. H. 1995. Development of human T and natural killer cells. *Blood.* 85: 2654-2670.

Storkus, W. J., Alexander, J., Payne, J. A., Dawson, J. R., and Cresswell, P. 1989. Reversal of natural killing susceptibility in target cells expressing transfected class I HLA genes. *Proceedings of the National Academy of Sciences of the United States of America.* 86: 2361-2364.

Sultmann, H., Murray, B. W., and Klein, J. 2000. Identification of seven genes in the Major Histocompatibility Complex class I region of the zebrafish. *Scandanavian Journal of Immunology.* 51: 577-585.

Takami, K., Zaleska-Rutczynska, Z., Figueroa, F., and Klein, J. 1997. Linkage of *LMP, TAP* and *RING3* with *Mhc* class I rather than class II genes in the zebrafish. *The Journal of Immunology.* 159: 6052-6060.

Takeuchi, H., Figueroa, F., O'hUigin, C., and Klein, J. 1995. Cloning and characterization of class I Mhc genes of the zebrafish, *Brachydanio rerio. Immunogenetics.* 42: 77-84.

Takeya, T. and Hanafusa, H. 1983. Structure and sequence of the cellular gene homologous to the RSV src gene and the machanism for generating the transforming virus. *Cell.* 32: 881-890.

Townsend, A., Elliott, T., Cerundolo, V., Foster, L., Barber, B., and Tse, A. 1990. Assembly of MHC class I molecules analyzed in vitro.[erratum appears in Cell 1990 Sep 21;62(6):following 1233]. *Cell.* 62: 285-95.

Townsend, A., Ohlen, C., Bastin, J., Ljunggren, H. G., Foster, L., and Karre, K. 1989. Association of class I major histocompatibility heavy and light chains induced by viral peptides.[comment]. *Nature.* 340: 443-8.

Townsend, A. R. M., Gotch, F. M., and Davey, J. 1985. Cytotoxic T cells recognize fragments of the influencza nucleoprotein. *Cell.* 42: 457-467.

Tragardh, L., Rask, L., Wiman, K., Fohlman, J., and Peterson, P. A. 1979. Amino acid sequence of an immunoglobulin-like HLA antigen heavy chain domain. *Proceedings of the National Academy of Sciences of the United States of America.* 76: 5839-42.

Trowsdale, J., Hanson, I., Mockridge, I., Beck, S., Townsend, A., and Kelly, A. 1990. Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters. *Nature.* 348: 741-4.

Tsuda, T. T., Tsuda, M., Naruse, T., Kawata, H., Ando, A., Shiina, T., Fukuda, M., Kurita, M., LeMaho, I., Kulski, J. K., and Inoko, H. 2001. Phylogenetic analysis of penguin (*Spheniscidae*) species based on sequence variation in MHC class II genes. *Immunogenetics.* 53: 712-716.

Tuttleman, J. S., Pugh, J. C., and Summers, J. W. 1986. In vitro experimental infection of primary duck hepatocyte cultures with duck hepatitis B virus. *Journal of Virology.* 58: 17-25.

Urban, M. K., O'Connell, A. P., and London, W. T. 1985. Sequence of events in natural infection of pekin duck embryos with duck hepatitis B virus. *Jouranl of Virology.* 55: 16-22.

van Erp, S. H. M., Dixon, B., Figueroa, F., Egberts, E., and Stet, R. J. M. 1996. Identification and characterization of a new major histocompatibility complex class I gene in carp (*Cyprinus carpio L.*). *Immunogenetics.* 44: 49-61.

van Tuinen, M. and Hedges, S. B. 2001. Calibration of avian molecular clocks. *Molecular Biology & Evolution.* 18: 206-13.

van Tuinen, M., Sibley, C. G., and Hedges, S. B. 2000. The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Molecular Biology & Evolution.* 17: 451-7.

Vega, M. A. and Strominger, J. L. 1989. Constitutive endocytosis of HLA class I antigens requires a specific portion of the intracytoplasmic tail that shares structural features with other endocytosed molecules. *Proceedings of the National Academy of Sciences of the United States of America.* 86: 2688-92.

Velarde, G., Ford, R. C., Rosenberg, M. F., and Powis, S. J. 2001. Three-dimensional structure of transporter associated with antigen processing (TAP) obtained by single Particle image analysis. *Journal of Biological Chemistry.* 276: 46054-63.

Verland, S., Simonsen, M., Gammeltoft, S., Allen, H., Flavell, R. A., and Olsson, L. 1989. Specific molecular interaction between the insulin receptor and a D product of MHC class I. *The Journal of Immunology.* 143: 945-951.

von Boehmer, H. 1994. Positive selection of lymphocytes. *Cell.* 76: 219-228.

von Schantz, R., Bensch, S., Grahn, M., Hasselquist, D., and Wittzell, H. 1999. Good genes, oxidative stress and condition-dependent sexual signals. *Proceedings of the Royal Society of London- Series B: Biological Sciences.* 266: 1-12.

von Schantz, T., Wittzell, H., Goransson, G., Grahn, M., and Persson, K. 1996. MHC genotype and male ornamentation: genetic evidence for the Hamilton-Zuk model. *Proceedings of the Royal Society of London- Series B: Biological Sciences.* 263: 265-271.

Vos, J. C., Reits, E. A., Wojcik-Jacobs, E., and Neefjes, J. 2000. Head-head/tail-tail relative orientation of the pore-forming domains of the heterodimeric ABC transporter TAP. *Current Biology.* 10: 1-7.

Vos, J. C., Spee, P., Momburg, F., and Neefjes, J. 1999. Membrane topology and dimérization of the two subunits of the transporter associated with antigen processing reveal a three-domain structure. *The Journal of Immunology.* 163: 6679-85.

Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO Journal.* 1: 945-51.

Wallny, H. and Rammensee, H. 1990. Identification of classical minor histocompatibility antigen as cell-derived peptide. *Nature.* 343.

Wang, C., Perera, T. V., Ford, H. L., and Dascher, C. C. 2003. Characterization of a divergent non-classical MHC class I gene in sharks. *Immunogenetics.* .

Watkins, D. I., Chen, Z. W., Hughes, A. L., Evans, M. G., Tedder, T. F., and Letvin, N. L. 1990. Evolution of the MHC class I genes of a New World primate from ancestral homologues of human non-classical genes. *Nature.* 346: 60-63.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. 1992. Evolution and ecology of Influenza A viruses. *Microbiological Reviews.* 56: 152-179.

Webster, R. G., Yakhno, M., Hinshaw, V. S., Bean, W. J., and Murti, K. G. 1978. Intestinal influenza: replication and characterization of influenza viruses in ducks. *Virology.* 84: 268-78.

Wedekind, C. and Furi, S. 1997. Body odour preferences in men and women: do they aim for specific MHC combinations or simply heterozygosity? *Proceedings of the Royal Society of London- Series B: Biological Sciences.* 264: 1471-1479.

Wedekind, C., Seebeck, T., Bettens, F., and Paepke, A. J. 1995. MHC-dependent mate preferences in humans. *Proceedings of the Royal Society of London- Series B: Biological Sciences.* 260: 245-249.

Wei, M. L. and Cresswell, P. 1992. HLA-A2 molecules in an antigen-processing mutant cell contain signal sequence-derived peptides. *Nature.* 356: 443-446.

Westerdahl, H., Wittzell, H., and von Schantz, T. 1999. Polymorphism and transcription of Mhc class I genes in a passerine bird, the great reed warbler. *Immunogenetics.* 49: 158-70.

Westerdahl, H., Wittzell, H., and von Schantz, T. 2000. Mhc diversity in two passerine birds: no evidence for a minimal essential Mhc. *Immunogenetics.* 52: 92-100.

Wiertz, E. J. H. J., Mukherjee, S., and Ploegh, H. L. 1997. Viruses use stealth technology to escape from the host immune system. *Molecular Medicine Today.* 3: 116-123.

Williams, D. B., Barber, B. H., Flavell, R. A., and Allen, H. 1989. Role of b 2-microglobulin in the intracellular transport and surface expression of murine class I histocompatibility molecules. *The Journal of Immunology.* 142: 2796-806.

Wittzell, H., von Schantz, T., Zoorob, R., and Auffray, C. 1994. Molecular characterization of three Mhc class II B haplotypes in the ring-necked pheasant. *Immunogenetics.* 39: 395-403.

Wittzell, H., von Schantz, T., Zoorob, R., and Auffray, C. 1995. Rfp-Y-like sequences assort independently of pheasant Mhc genes. *Immunogenetics.* 42: 68-71.

Wu, J., Cherwinski, H., Spies, T., Phillips, J. H., and Lanier, L. L. 2000. DAP10 and DAP12 form distinct, but functionally cooperative, receptor complexes in natural killer cells. *Journal of Experimental Medicine.* 192: 1059-1067.

Wu, J., Song, Y., Bakker, A. B. H., Bauer, S., Spies, T., Lanier, L. L., and Phillips, J. H. 1999. An activating Immunoreceptor complex formed by NKG2D and DAP10. *Science.* 285: 730-732.

Yamazaki, K., Beauchamp, G. K., Kupniewski, D., Bard, J., Thomas, L., and Boyse, E. A. 1988. Familial imprinting determines H-2 selective mating preferences. *Science.* 240: 1331-1332.

Yang, Y., Waters, J. B., Fruh, K., and Peterson, P. A. 1992. Proteasomes are regulated by interferon g: Implications for antigen processing. *Proceedings of the National Academy of Sciences of the United States of America.* 89: 4928-4932.

Yokoyama, W. M. 1999. Natural Killer Cell. *In* W. E. Paul (ed.): *Fundamental Immunology*, Lippincott Raven, Philadelphia.

Younger, R. M., Amadou, C., Bethel, G., Ehlers, A., Lindahl, K. F., Forbes, S., Horton, R., Milne, S., Numgall, A. J., Trowsdale, J., Volz, A., Ziegles, A., and Beck, S. 2001. Characterization of clustered MHC-linked olfactory receptor genes in human and mouse. *Genome Research.* 11: 519-530.

Zinkernagel, R. M., Callahan, G. N., Althage, A., Cooper, S., Klein, P. A., and Klein, J. 1978. On the thymus in the differentiation of "H-2 self-recognition" by T cells: evidence for dual recognition? *The Journal of Experimental Medicine.* 147: 882-896.

Zinkernagel, R. M. and Doherty, P. H. 1974. Restriction of *in vitro* T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature.* 248: 701-702.

Zoorob, R., Bernot, A., Renoir, D. M., Choukri, F., and Auffray, C. 1993. Chicken major histocompatibility complex class II B genes: analysis of interallelic and interlocus sequence variance. *European Journal of Immunology.* 23: 1139-1145.

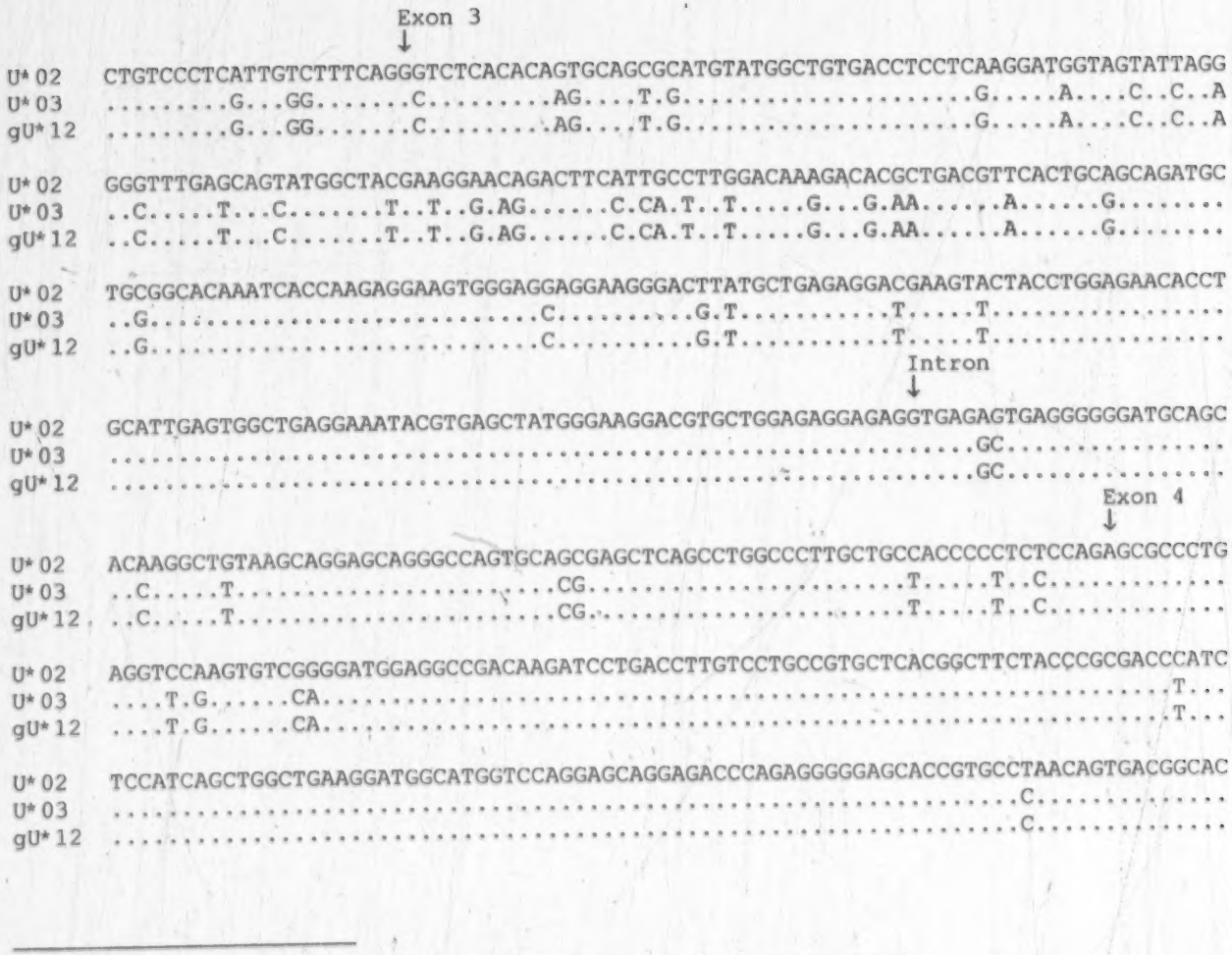
1 APPENDIX A: NUCLEOTIDE SEQUENCES OF MHC CLASS I AND TAP2 AMPLIFIED FROM GENOMIC DNA OF D26

```
                Exon 1                                                                                                  Intron
                  ↓                                                                                                       ↓
U*02    ATGGGACCGGGACGGGCGGGGGCCCTGGGCCTCGTCCTGGGGCTGCTGCTGGGGGTCCTGGGCGGAGCAACGAGTGGTGA
U*03    ..............................T.G.G.....................................G....C.....
gU*12(2)-------------------------------------------------------------*.........G....C.....

U*02    GTGTGAGTGGGGCCGGGCCCTCCTGGCACCGGGACCCCCCGGGCCTCCCCTCTCCTCCCCTTCCCCTCTCGGCCCCCCCT
U*03    ..A........T.-.....................................................A.............-.
gU*12   ..A........T.-.....................................................A.......N.....-.
                                                                  Exon 2
                                                                    ↓
U*02    GGGCTCGGCCTCCCCTTGGTGCTGCCCCTCTCAGCCGCCCCTGTTCCCGCAGAGCCCCACTCCCTGCGCTATTTCTACAC
U*03    ................C..GA.....................................................................
gU*12   ................C..GA.....................................T.....................

U*02    CGCGGTGTCAGAACCGAGCCCGGGGCTGCCTCAGTTCGTGGGCGTGGGGTATGTGGATGGGGAGGCCTTCGTGCGCTATG
U*03    ........G..C...........G....G....A.....CT.....A..C...........A......ACATA.....
gU*12   ................................................................................

U*02    ACAGCGAGACCCACAGGATGGATTCCATGGTGGACTGGACGTCGGCCATTGATGACCAGCAGTACTGGGAGTGGAACACC
U*03    ....T......AGG....C...GC..CG...........TTG......A.ATG..T..............CAT.GTG..T
gU*12   ...............................................C................................
                                                                                                        Intron
                                                                                                          ↓
U*02    CAGAACTTTCAGAATGATGAGAAGATTTTCCGCGTGAACCTGGACACGCTGCGGGAGCGCTACAACCAGAGCAGGGGTGA
U*03    G.A.T...C.....CAC....C......A....A.C...........A................................
gU*12   ................................................................................
```

* Figure A 1: MHC class I genomic PCR fragment amplified from exon 1 to exon 6 (continued).

```
U*02    GAGCGGGCTGGGCCACAGCTCTGCAGGCTGTGGGAGCGGACGCTGTGGGAGGGGTGGGGGGGTCCCCAGCCCCACTGAGG
U*03    ........................C............A.GT.....................................
gU*12   ..............................................................................

U*02    CCTGGCCCTGCCCCCACACTACTGCCCCAGCCCTGGCTGTGCTGTTTGGTCCCTGTGTGGTGTTGCTGTCACAGGGCTCC
U*03    .........T...G...G.....................................................................
gU*12   .........T...G...G.....................................................................

U*02    CAGCTGCCCTGTCCCAGCACCACGATGCTCCTGGTGCCTTGTCCCAGACCCAGAGGGGTGCTTCCAGACAGAACCCAGCC
U*03    ..............................................................................
gU*12   ..............................................................................

U*02    CCTTGCCACCCCCACGCTGCTCAGAACCCCCTGTCCCGGAGCTGGGGAGCTCCAGGCCACCCCATCCCAGCCCTTCCACA
U*03    ............G.T.............A....C...A...................T......C.......T........
gU*12   ............G.T.............A....C...A...................T......C.......T........

U*02    TACCCACAGCAGTCAGGACCCCCAGAGCCAGAGGGACTCCCGCCCCTGTTGTCCCAGCCCCTTGCTTTTCCCATGGTGCT
U*03    .CTAG.....................A........G..T.T.....CA.....................C.....C.....
gU*12   .CTAG.....................A........G..T.T.....CA.....................C.....C.....

U*02    TCAGGCCTCCACACCAAAACCGGAGTGATTCCCCCCACCCCATCTCAGCTCCTGGTCACCTCACAGCATTGACAACTCCG
U*03    .TT...TC....C........T...G.....................................................C.T.
gU*12   .TT...TC....C........T...G.....................................................C.T.

U*02    TACAAGGCCCTTGCTGCCCCAACATACTGTGGTGCTGGGCCCGAGCCCAGTCTTTCAGGGATGCAGCACTCTGGGACAGG
U*03    .......T...C...C......AG.................C.................A.....................
gU*12   .......T...C...C......AG.................C.................A.....................

U*02    GCAGTGTGTGGCACCTGCATGCGCTGCCAGTGGGGTCTGCAGCTCATTAAAGGGCCAGGACTGTGTCTGCCCCTCACCCC
U*03    ..............C......A...A.............C.......................G..........A...
gU*12   ..............C......A...A.............C.......................G..........A...
```

*Figure A 1: MHC class I genomic PCR fragment amplified from exon 1 to exon 6 (continued).



* Figure A 1: MHC class I genomic PCR fragment amplified from exon 1 to exon 6 (continued).

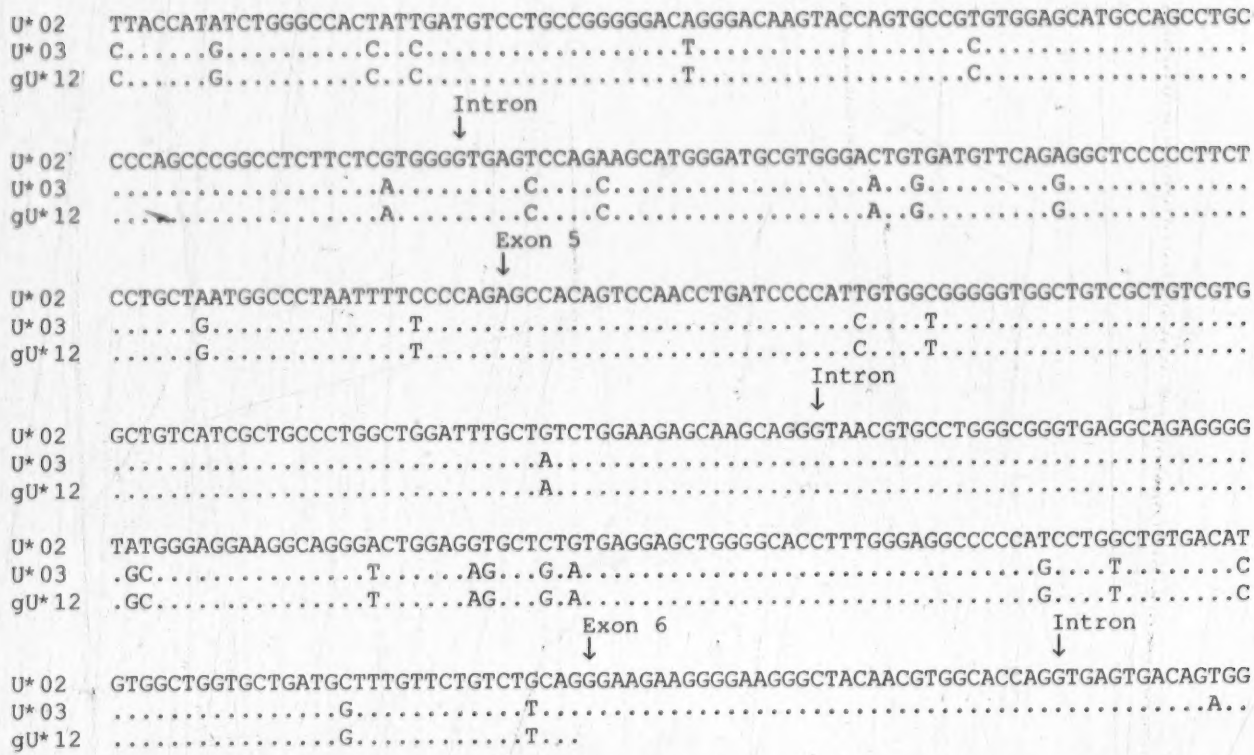
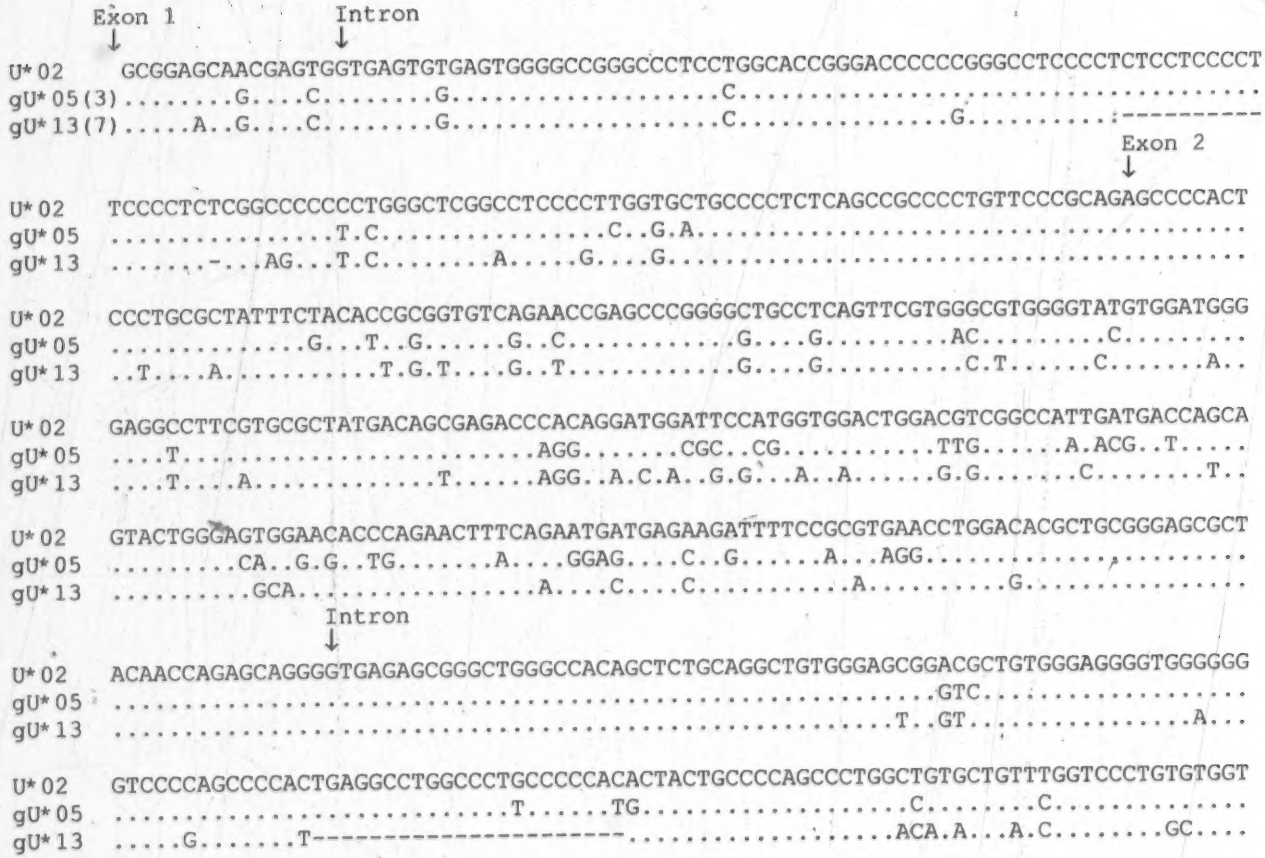


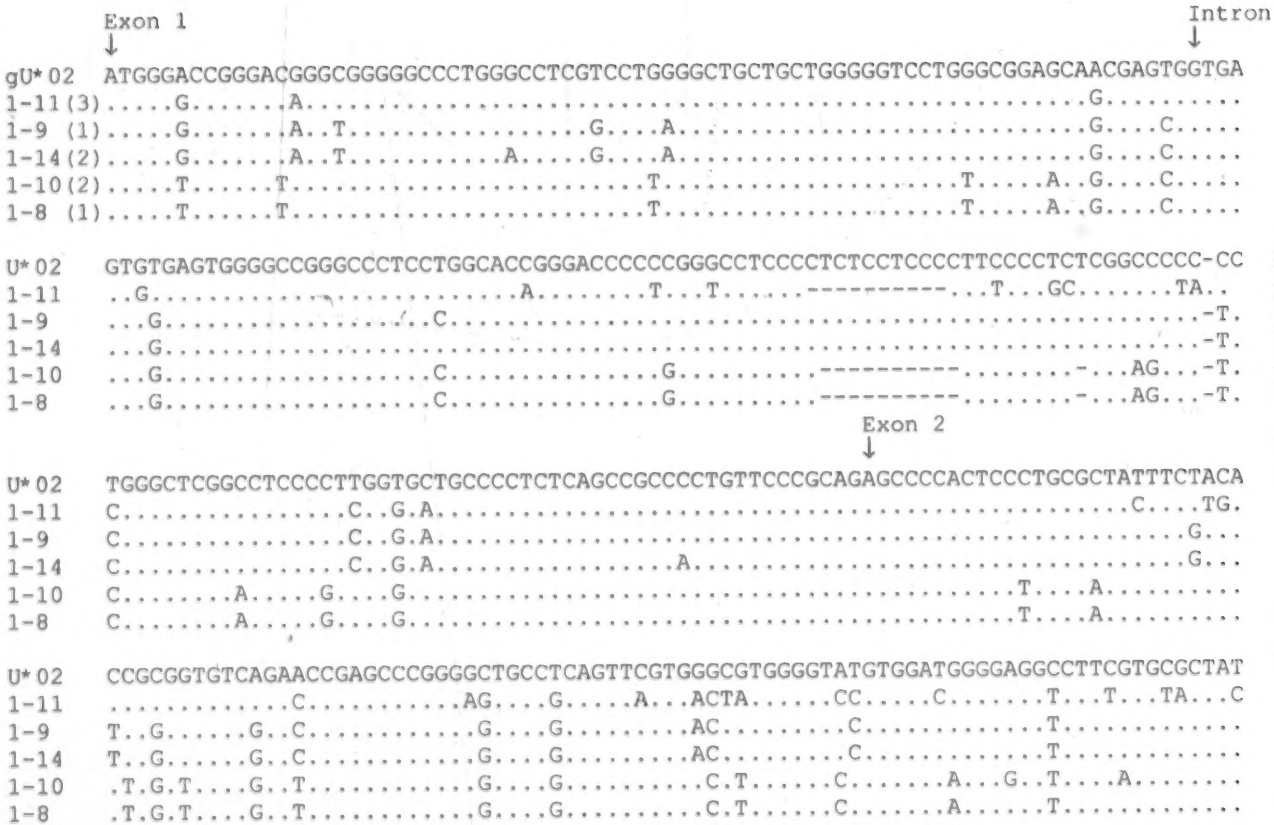
Figure A 1: MHC class I genomic PCR fragment amplified from exon 1 to exon 6. Primers D26SP2 and D26CY1were used to amplify MHC class I products from D26. The PCR products were cloned and sequenced. One sequence, gU*12, was identified that shared high nucleotide identity with U*02 in exon 2, and high nucleotide identity with U*03 through the rest of the sequence. Clones containing coding region sequences identical to the cDNA clones U*02 and U*03 were also identified. The sequence is aligned to the genomic sequence of AnplU*02 and U*03. Numbers in brackets represent the number of clones identified for each sequence. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02.



* Figure A 2: MHC class I genomic PCR fragment amplified from exon 1 to intron 2 (continued).

```
U*02    GTTGCTGTCACAGGGCTCCCAGCTGCCCTGTCCCAGCACCACGATGCTCCTGGTGCCTTGTCCCAGACCCAGAGGGGTGC
gU*05   ..........TG..............G.........C..........G..............C..............
gU*13   ..........TG..............G.........C............A............A.C......TA..A...

U*02    TTCCAGACAGAACCCAGCCCCTTGCCACCCCCACGCTGCTCAGAACCCCCTGTCCCGGAGCTGGGGAGCTCCAGGCCACC
gU*05   .........T.....................T.T...........T.....A....A.........G...........A.
gU*13   .........C.........T.......TG.....A.AT........A.........A....A....G........T....

U*02    CCATCCCAGCCCTTCCACATACCCACAGCAGT
gU*05   ..C..................C...........
gU*13   .........T...........C...........
```

Figure A 2: MHC class I genomic PCR fragment amplified from exon 1 to intron 2. The primer D26SP2 reversed primed in intron 2 of some MHC class I sequences from D26. The PCR products were cloned and sequenced. Two sequences were generated by reverse priming of the forward primer were identified and aligned to the genomic sequence of *Anpl*U*02. Numbers in brackets represent the number of clones identified for each sequence. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02. Dashes are included to maximize identity of the alignment.



* Figure A3: MHC class I genomic PCR fragment amplified from the 5'untranslated region to exon 2 (continued).

```
U*02    GACAGCGAGACCCACAGGATGGATTCCATGGTGGACTGGACGTCGGCCATTGATGACCAGCAGTACTGGGAGTGGAACAC
1-11    .....T.....GAGG...A...GGT.CG..C........TTG.....CA.ACG..T.............CA..C.G..
1-9     ............AGG.......CGC..CG...........TTG......A.ACG..T.............CA..G.G..
1-14    ............AGG........AC..CG...........TTG......A.ACG..T.............CA..G.G..
1-10    .....T......AGG..A.C.A..G.G...A..A......G.G.......C........T...........GCA.....
1-8     ............AGG........AC..CG...........TTG......A.ACG..T.............CA..G.G..

U*02    CCAGAACTTTCAGAATGATGAGAAGATTTTCCGCGT-GAACCTGGACACGCTG
1-11    TG...CTGCA.G..GGA.....C.............-.GG..............
1-9     TG.......A....GGAG....C..G......A...A-GG.............
1-14    TG.......A....GGAG....C..G......A...A-G.A............
1-10    ............A....C....C...........A.-.........G......
1-8     TG.......A....GGAG....C..G......A...A-GG.............
```

Figure A 3: MHC class I genomic PCR fragment amplified from the 5'untranslated region to exon 2. Primers 5'UTF2 and D26E2R1 were used to amplify MHC class I genomic sequences from D26. The products were cloned and sequenced. Sequence 1-11 matches U*04, sequence 1-9 matches U*05, and three unique sequences, 1-14, 1-10, and 1-8, were identified and aligned with the genomic sequence of *Anpl*U*02. Numbers in brackets represent the number of clones identified for each sequence. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02. Dashes are included to maximize identity of the alignment.

```
        Exon 2
          ↓
U*02    TTCTACACCGCGGTGTCAGAACCGAGCCCGGGGCTGCCTCAGTTCGTGGGCGTGGGGTATGTGGATGGGGAGGCCTTCGT
2-4(3)  ....TG.............C...........AG....G.....A...ACTA......CC.....C.......T...T..
2-8(6)  ...G...T..G......G..C...........G....G.........AC.........C.............T......
2-1(6)  ...G...T..G......G..C...........G....G.........AC.........C.............T......

U*02    GCGCTATGACAGCGAGACCCACAGGATGGATTCCATGGTGGACTGGACGTCGGCCATTGATGACCAGCAGTACTGGGAGT
2-4     .TA...C.....T.....GAGG....A...GGT.CG..C.......TTG.....CA.ACG..T.............CA
2-8     ...................AGG.......CGC..CG...........TTG......A.ACG..T.............CA
2-1     ...................AGG........AC..CG...........TTG......A.ACG..T.............CA

U*02    GGAACACCCAGAACTTTCAGAATGATGAGAAGATTTTCCGCGTGAACCTGGACACGCTGCGGGAGCGCTACAACCAGAGC
2-4     ..C.G..TG...CTGCA.G..GGA.....C.............GG..................................
2-1     ..G.G..TG.......A....GGAG....C..G......A...AGG..................................
2-8     ..G.G..TG.......A....GGAG....C..G......A...AGG..................................
            Intron
            ↓
U*02    AGGGGTGAGAGCGGGCTGGGCCACAGCTCTGCAGGCTGTGGGAGCGGACGCTGTGGGAGGGGTGGGGGGGTCCCCAGCCC
2-4     ..................A.............GG........T.......A.CA...........................
2-8     ..........................................................GT.....................
2-1     ..................A...........................GT.....................................

U*02    CACTGAGGCCTGGCCCTGCCCCCACACTACTGCCCCAGCCCTGGCTGTGCTGTTTGGTCCCTGTGTGGTGTTGCTGTCAC
2-4     ..A..............T....A.-------------.........CT.............T...................
2-8     ..................T......TG.....................C........C.......................T
2-1     ..A..............T......TG.....................C........C.......................T

U*02    AGGGCTCCCAGCTGCCCTGTCCCAGCACCACGATGCTCCTGGTGCCTTGTCCCAGACCCAGAGGGGTGCTTCCAGACAGA
2-4     G.........................C...T...................G...C.......................T.
2-8     G..............G.........C..........G..............C.......................T.
2-1     G...........CA............C...........................C.......................T.
```

* Figure A4: MHC class I genomic PCR fragment amplified from exon 2 to exon 3 (continued).

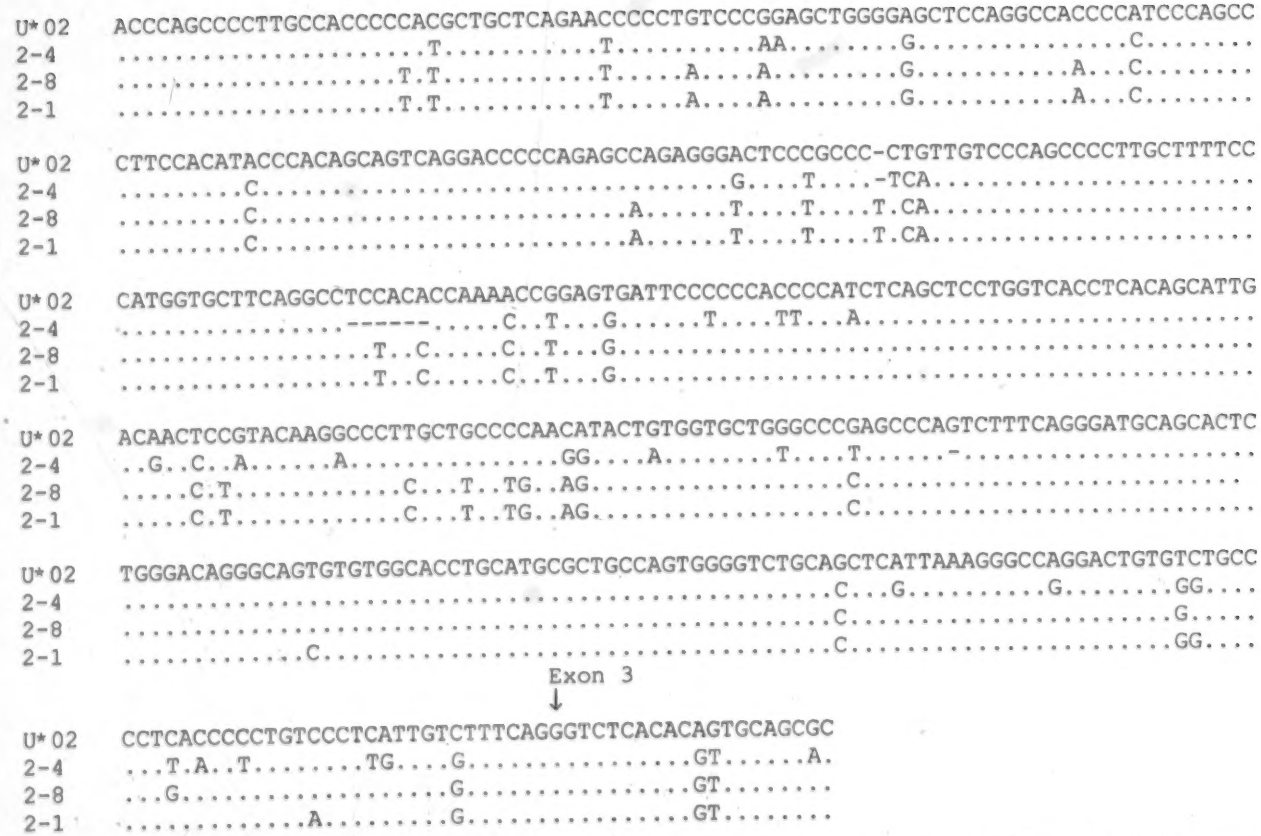


Figure A 4: MHC class I genomic PCR fragment amplified from exon 2 to exon 3. Primers D26E2F1 and D26E3R1were used to amplify MHC class I products from D26. The PCR products were cloned and sequenced. Three unique sequences, 2-4, 2-8 and 2-1, were identified and aligned to the genomic sequence of *Anpl*U*02. Numbers in brackets represent the number of clones identified for each sequence. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02. Dashes are included to maximize identity of the alignment.

```
        Exon 3
          ↓
U*02    AGGATGGTAGTATTAGGGGGTTTGAGCAGTATGGCTACGAAGGAACAGACTTCATTGCCTTGGACAAAGACACGCTGACG
3-12(5)........................T...........T.....G........TC..........TG......T.......
3-14(1)........................T...........T.....G........TC..........TG......T.......
3-1(1) ........................T...........T.....G........TC..........TG......T.......

U*02    TTCACTGCAGCAGATGCTGCGGCACAAATCACCAAGAGGAAGTGGGAGGAGGAAGGGACTTATGCTGAGAGGACGAAGTA
3-12    .A.........G..C.....A..T.........................C...........G........C.C.G......
3-14    .A.........G..C.....A..T.........................C...........G........C.C.G......
3-1     .A.........G..C.....A..T.........................C...........G........C.C.G......
                                                                              Intron
                                                                                ↓
U*02    CTACCTGGAGAACACCTGCATTGAGTGGCTGAGGAAATACGTGAGCTATGGGAAGGACGTGCTGGAGAGGAGAGGTGAGA
3-12    ........................................A.......................................
3-14    ........................................A.......................................
3-1     ........................................A......T..........A.....................G
4-5(1)  -------------------------.............C.T.....................................G
4-6(4)  -------------------------.............A......T................................G
4-1(1)  -------------------------.............A................T.....................G

U*02    GTGAGGGGGGATGCAGCACAAGGCTGTAAGCAGGAGCAGGGCCAGTGCAGCGAGCTCAGCCTGGCCCTTGCTGCCACCCC
3-12    ...................C.............................G........G...................
3-14    ...................C.............................G........G...................
3-1     C......T...........C.............................G............................
4-5     C...C..............C......G.C.......T............G........G...................A
4-6     C..................C.............................G............................
4-1     C...............T..C.....T..T.................A...G...T.........................
```

* Figure A5: MHC class I genomic PCR fragment amplified from exon 3 to exon 4 (continued).

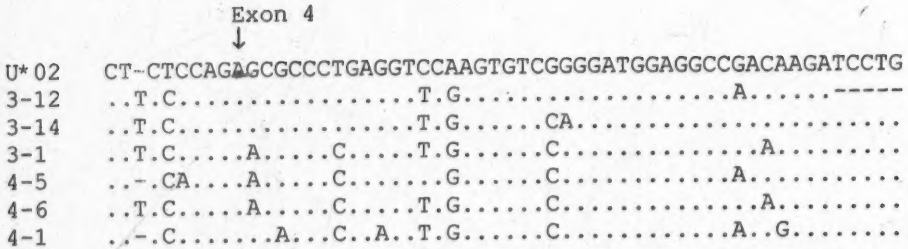
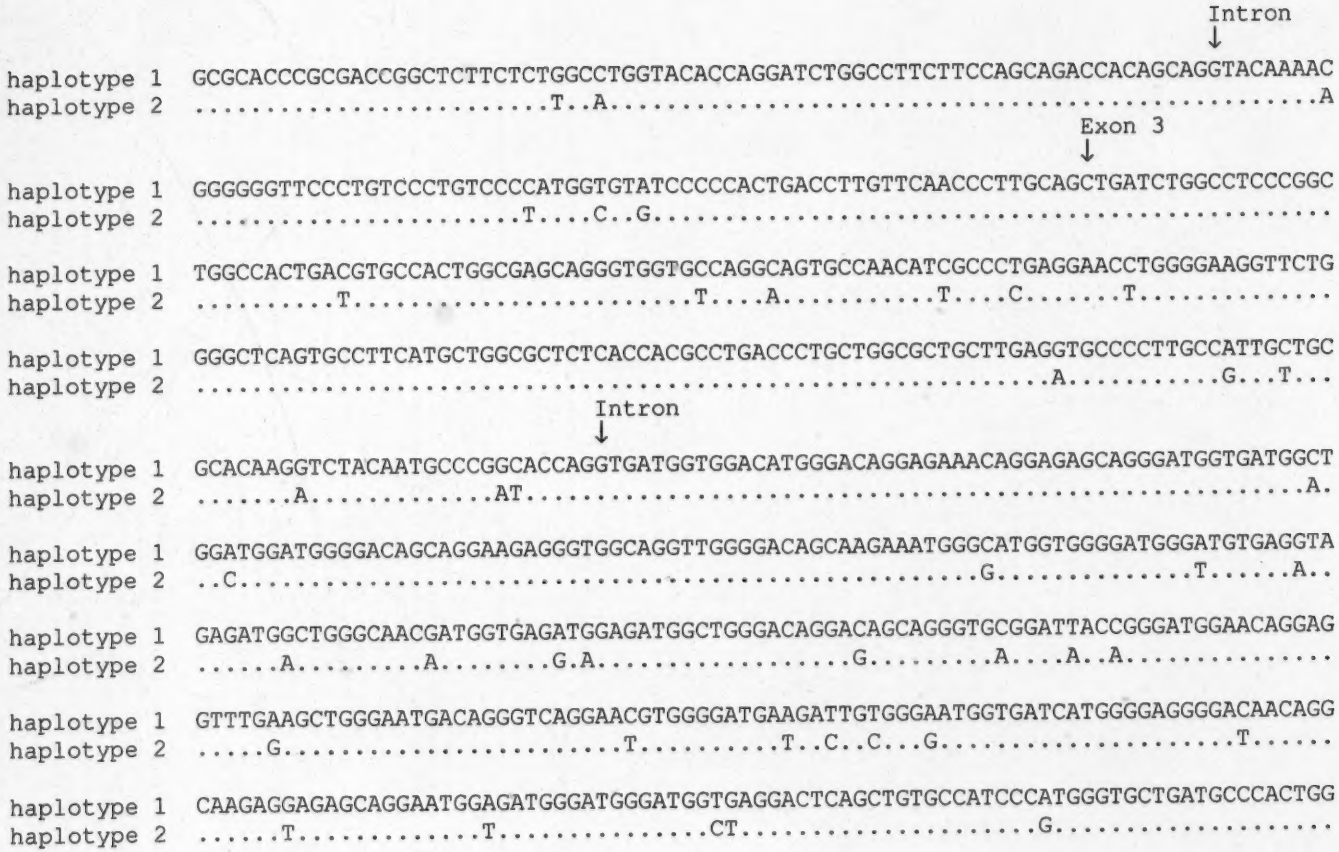
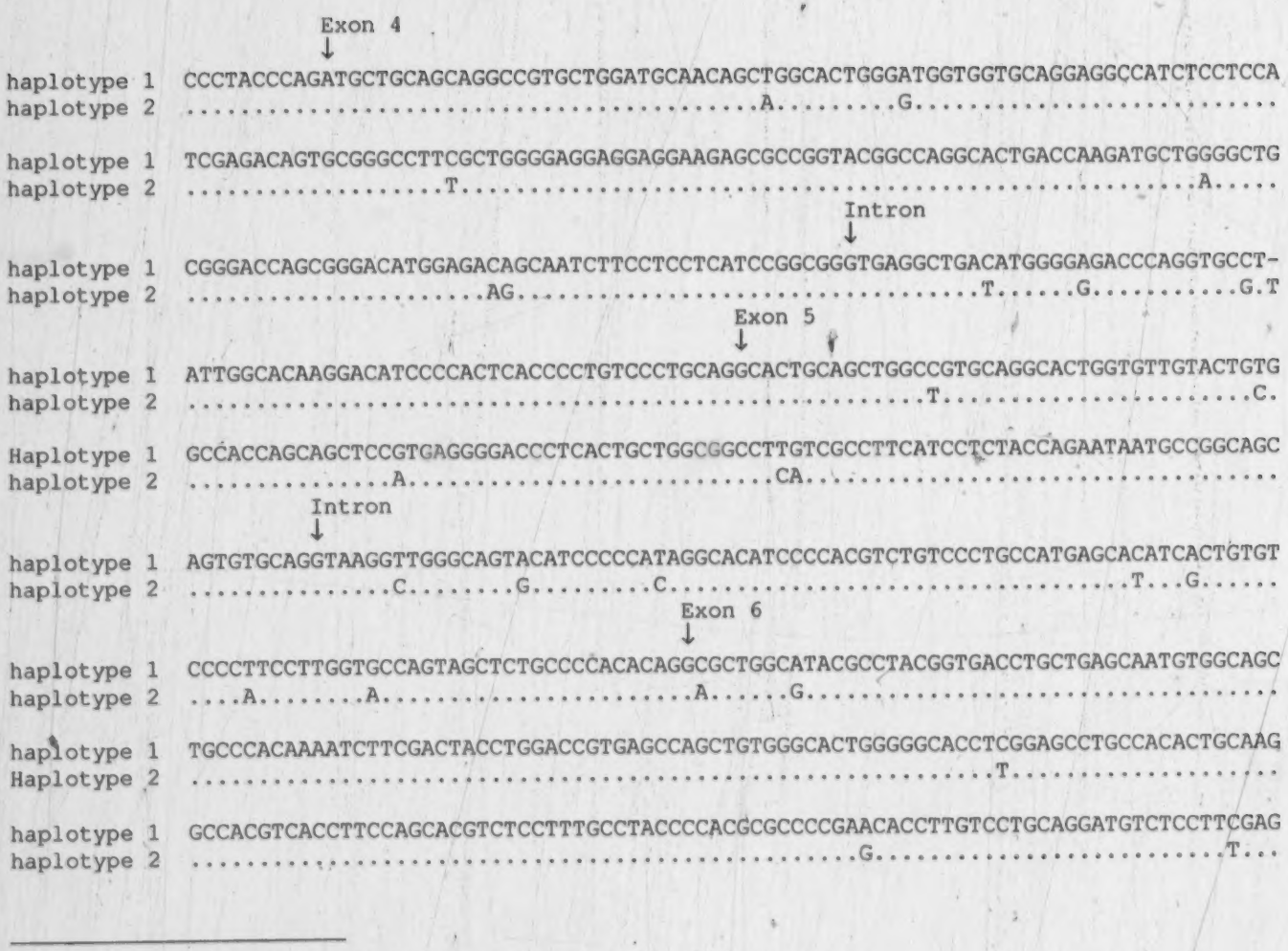


Figure A 5: MHC class I genomic PCR fragment amplified from exon 3 to exon 4. Primers D26E3F1 or D26E3F2 and D26E4R2 were used to amplify MHC class I products from D26. The PCR products were cloned and sequenced. Six unique sequences were identified and aligned to the genomic sequence of *Anpl*U*02. Numbers in brackets represent the number of clones identified for each sequence. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02. Dashes are included to maximize identity of the alignment.

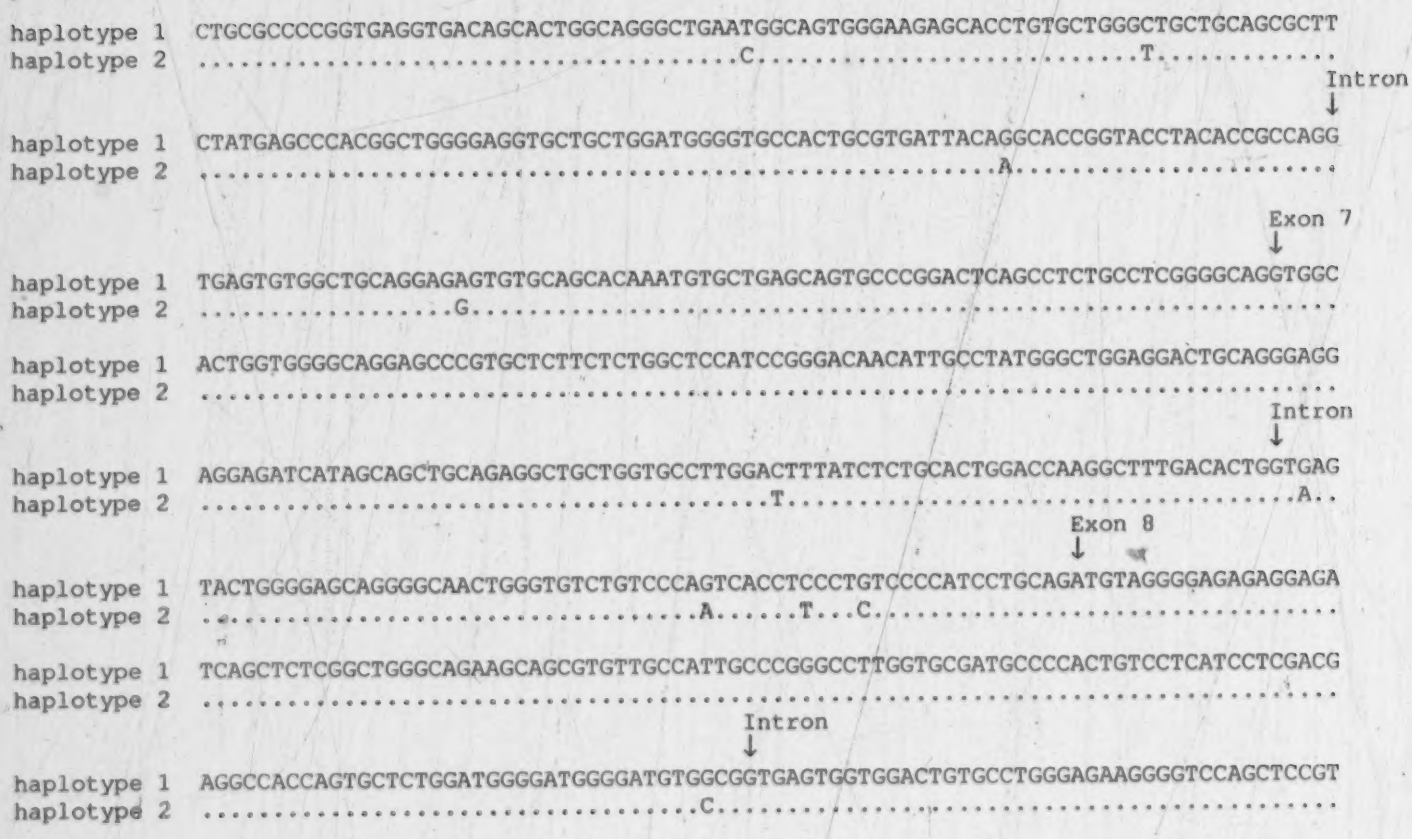
* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).



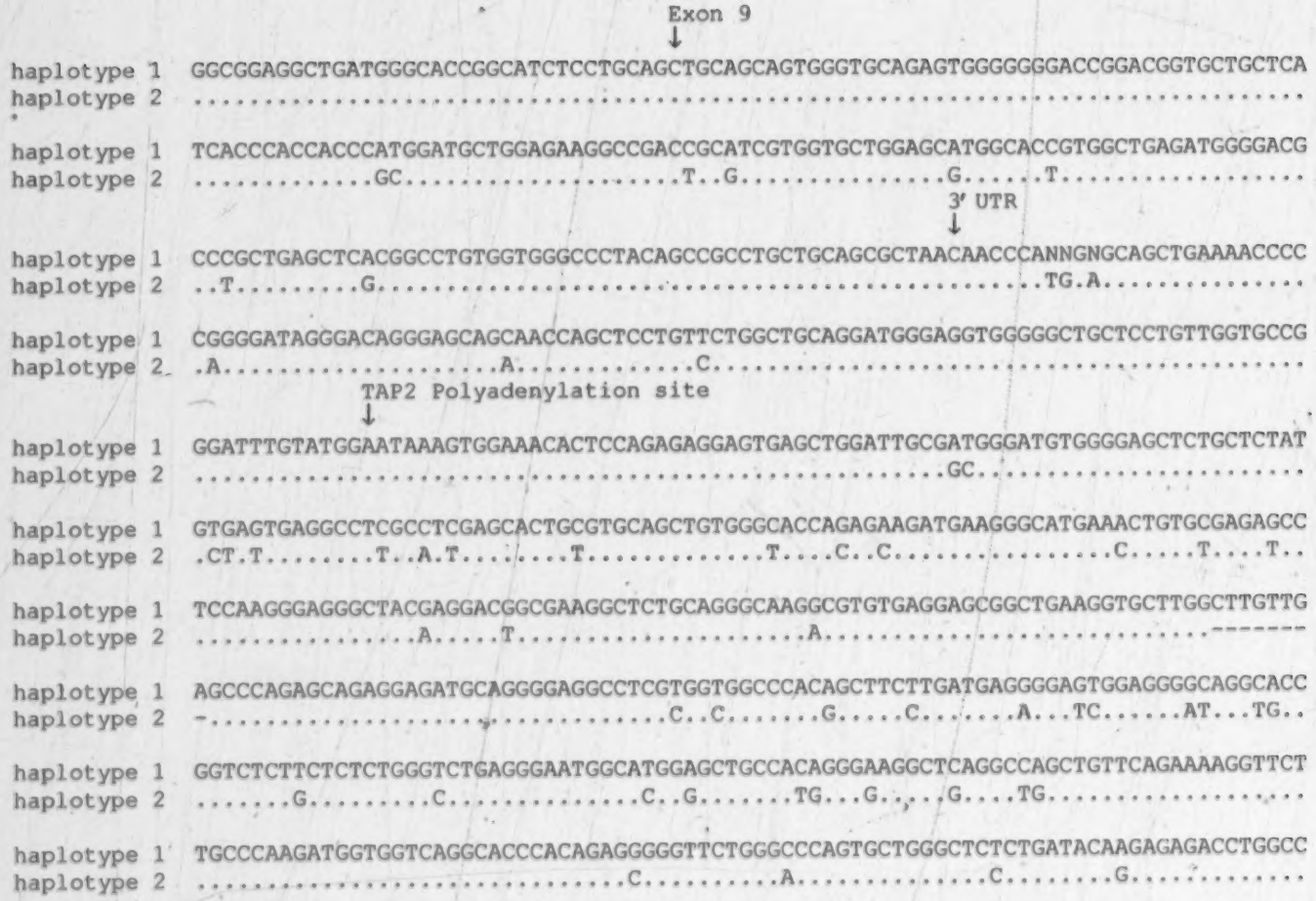
* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).



* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).



* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).

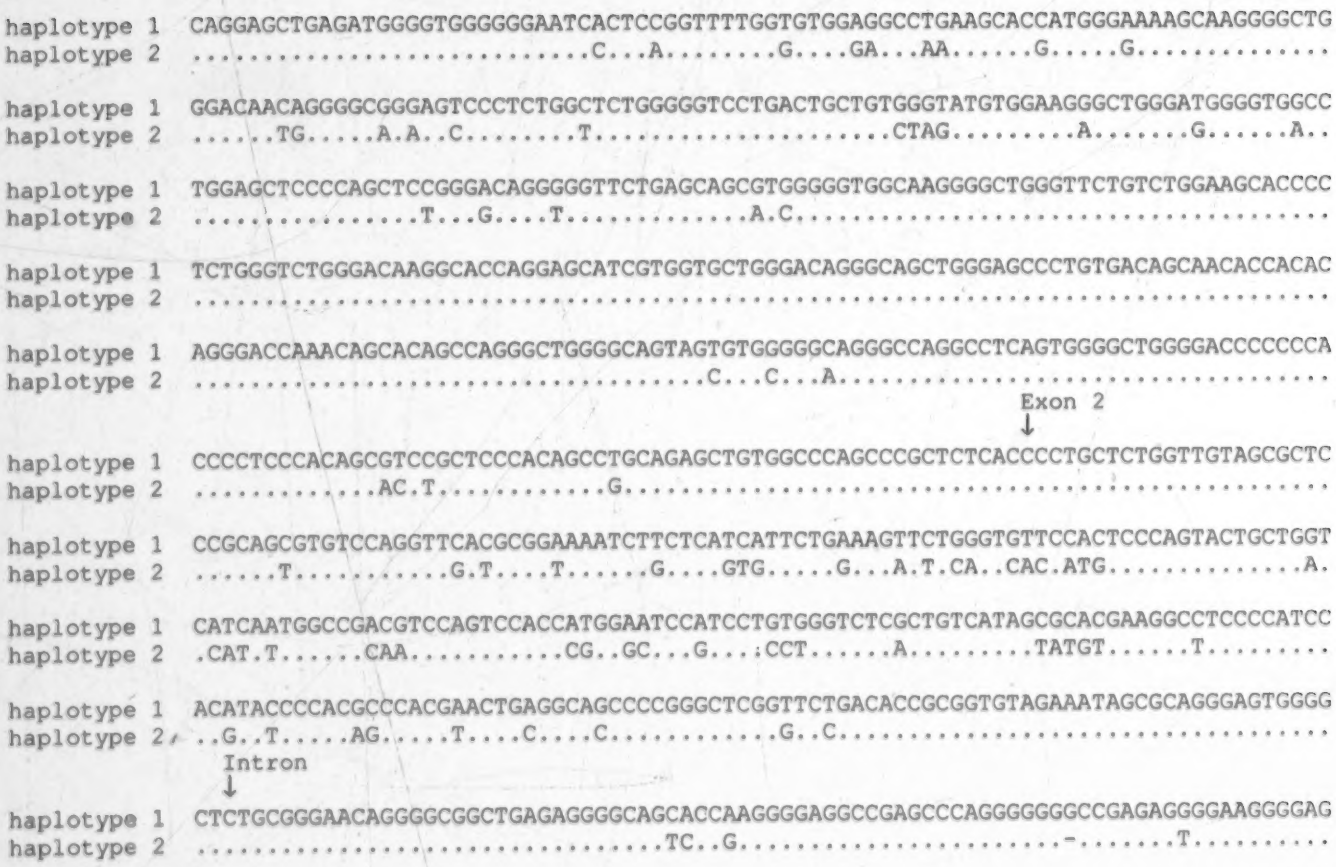


* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).

* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).

* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).

* Figure A6: Nucleotide sequence of haplotypes 1 and 2 (continued).



* Figure A 6: Nucleotide sequence of haplotypes 1 and 2 (continued).

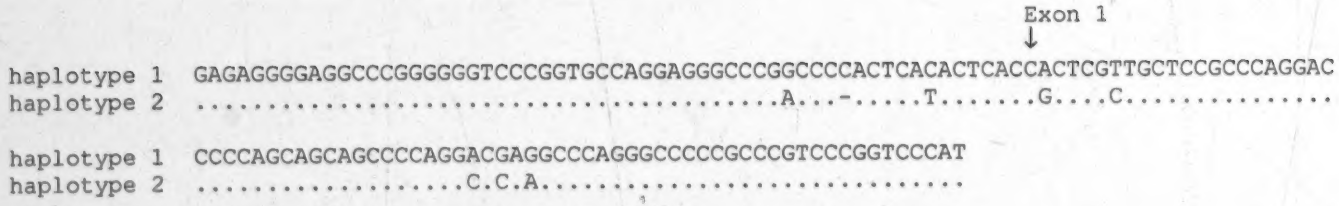
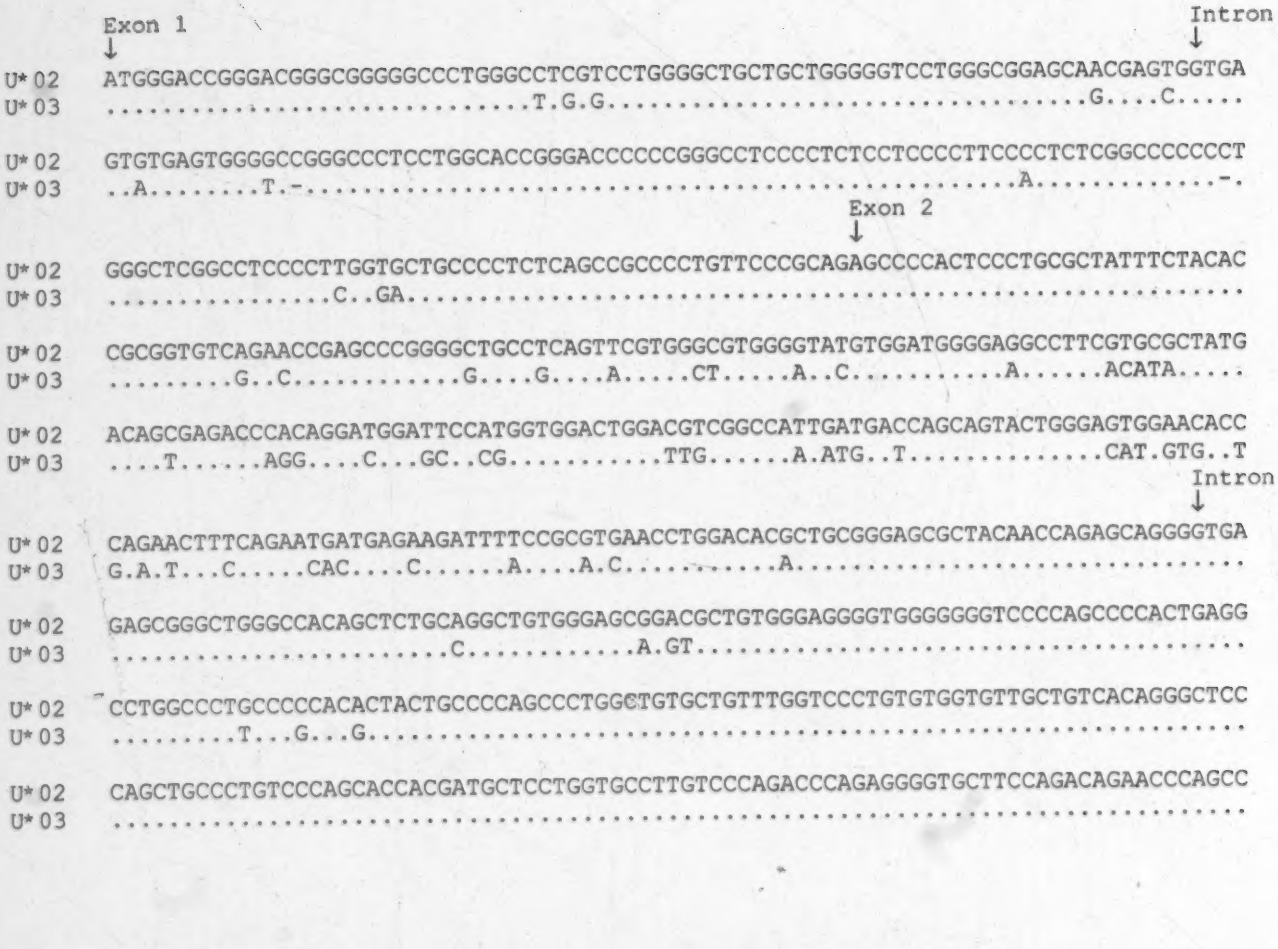


Figure A 6: Nucleotide sequence of haplotypes 1 and 2. The haplotypes identified span the TAP2 locus, TAP2/MHC class I intergenic region, and the dominantly expressed MHC class I locus of duck. Haplotype 1: TAP2B and U*02; Haplotype 2: TAP2A and U*03. PCR amplification and sequencing of D26 genomic DNA of TAP2, MHC class I, and from exon 5 of MHC class I to exon 7 of TAP2, allowed this contig of the major locus to be generated. Arrows mark the beginning of introns and exons. Dots indicate identity with haplotype 1. Dashes are included to maximize identity of the alignment.



* Figure A7: Genomic nucleotide alignment of the dominantly expressed MHC class I alleles (continued).

```
U*02   CCTTGCCACCCCCACGCTGCTCAGAACCCCCTGTCCCGGAGCTGGGGAGCTCCAGGCCACCCCATCCCAGCCCTTCCACA
U*03   ...........G.T...........A....C...A.................T......C.......T........

U*02  TACCCACAGCAGTCAGGACCCCCAGAGCCAGAGGGACTCCCGCCCCTGTTGTCCCAGCCCCTTGCTTTTCCCATGGTGCT
U*03  .CTAG.....................A........G..T.T.....CA...................C.....C.....

U*02   TCAGGCCTCCACACCAAAACCGGAGTGATTCCCCCCACCCCATCTCAGCTCCTGGTCACCTCACAGCATTGACAACTCCG
U*03   .TT...TC....C........T...G.....................................................C.T.

U*02   TACAAGGCCCTTGCTGCCCCAACATACTGTGGTGCTGGGCCCGAGCCCAGTCTTTCAGGGATGCAGCACTCTGGGACAGG
U*03   .......T...C...C......AG..................C..............A.....................

U*02   GCAGTGTGTGGCACCTGCATGCGCTGCCAGTGGGGTCTGCAGCTCATTAAAGGGCCAGGACTGTGTCTGCCCCTCACCCC
U*03   ..............C......A...A.............C.....................G..........A...
                              Exon 3
                              ↓
U*02   CTGTCCCTCATTGTCTTTCAGGGTCTCACACAGTGCAGCGCATGTATGGCTGTGACCTCCTCAAGGATGGTAGTATTAGG
U*03   .........G...GG.......C.........AG....T.G.....................G.....A....C..C..A

U*02   GGGTTTGAGCAGTATGGCTACGAAGGAACAGACTTCATTGCCTTGGACAAAGACACGCTGACGTTCACTGCAGCAGATGC
U*03   ..C.....T...C.......T..T..G.AG......C.CA.T..T.....G...G.AA......A......G........

U*02   TGCGGCACAAATCACCAAGAGGAAGTGGGAGGAGGAAGGGACTTATGCTGAGAGGACGAAGTACTACCTGGAGAACACCT
U*03   ..G.............................C...........G.T...........T.....T.................
                                                              Intron
                                                              ↓
U*02   GCATTGAGTGGCTGAGGAAATACGTGAGCTATGGGAAGGACGTGCTGGAGAGGAGAGGTGAGAGTGAGGGGGGATGCAGC
U*03   ..............................................................GC................
                                                                      Exon 4
                                                                      ↓
U*02   ACAAGGCTGTAAGCAGGAGCAGGGCCAGTGCAGCGAGCTCAGCCTGGCCCTTGCTGCCACCCCCTCTCCAGAGCGCCCTG
U*03   ..C.....T.......................CG.....................T.....T..C.............
```

* Figure A7: Genomic nucleotide alignment of the dominantly expressed MHC class I alleles (continued).

```
U*02  AGGTCCAAGTGTCGGGGATGGAGGCCGACAAGATCCTGACCTTGTCCTGCCGTGCTCACGGCTTCTACCCGCGACCCATC
U*03  ....T.G......CA.....................................................T...

U*02  TCCATCAGCTGGCTGAAGGATGGCATGGTCCAGGAGCAGGAGACCCAGAGGGGGAGCACCGTGCCTAACAGTGACGGCAC
U*03  ..........................................................C.............

U*02  TTACCATATCTGGGCCACTATTGATGTCCTGCCGGGGGACAGGGACAAGTACCAGTGCCGTGTGGAGCATGCCAGCCTGC
U*03  C......G..........C..C.................T.................C...................
                          Intron
                            ↓
U*02  CCCAGCCCGGCCTCTTCTCGTGGGGTGAGTCCAGAAGCATGGGATGCGTGGGACTGTGATGTTCAGAGGCTCCCCCTTCT
U*03  ...................A.........C....C.................A..G.........G.............
                              Exon 5
                                ↓
U*02  CCTGCTAATGGCCCTAATTTTCCCCAGAGCCACAGTCCAACCTGATCCCCATTGTGGCGGGGGTGGCTGTCGCTGTCGTG
U*03  ......G.............T.............................C....T.......................
                                                  Intron
                                                    ↓
U*02  GCTGTCATCGCTGCCCTGGCTGGATTTGCTGTCTGGAAGAGCAAGCAGGGTAACGTGCCTGGGCGGGTGAGGCAGAGGGG
U*03  ............................A.....................................................

U*02  TATGGGAGGAAGGCAGGGACTGGAGGTGCTCTGTGAGGAGCTGGGGCACCTTTGGGAGGCCCCCATCCTGGCTGTGACAT
U*03  .GC..............T......AG...G.A.................................G....T........C
                                      Exon 6                                Intron
                                        ↓                                     ↓
U*02  GTGGCTGGTGCTGATGCTTTGTTCTGTCTGCAGGGAAGAAGGGGAAGGGCTACAACGTGGCACCAGGTGAGTGACAGTGG
U*03  ..............G............T.....................................................A..

U*02  CAGCGATGTCTCCAGCATCTGCCCATCACTGGGGCAGTCTTGGGGCTCTGTTTTTTCTTGGGGCTCCCAGCACAGCGGTG
U*03  ....A.....C.....G...C.............................-.GG.....C--------------------
```

* Figure A7: Genomic nucleotide alignment of the dominantly expressed MHC class I alleles (continued).

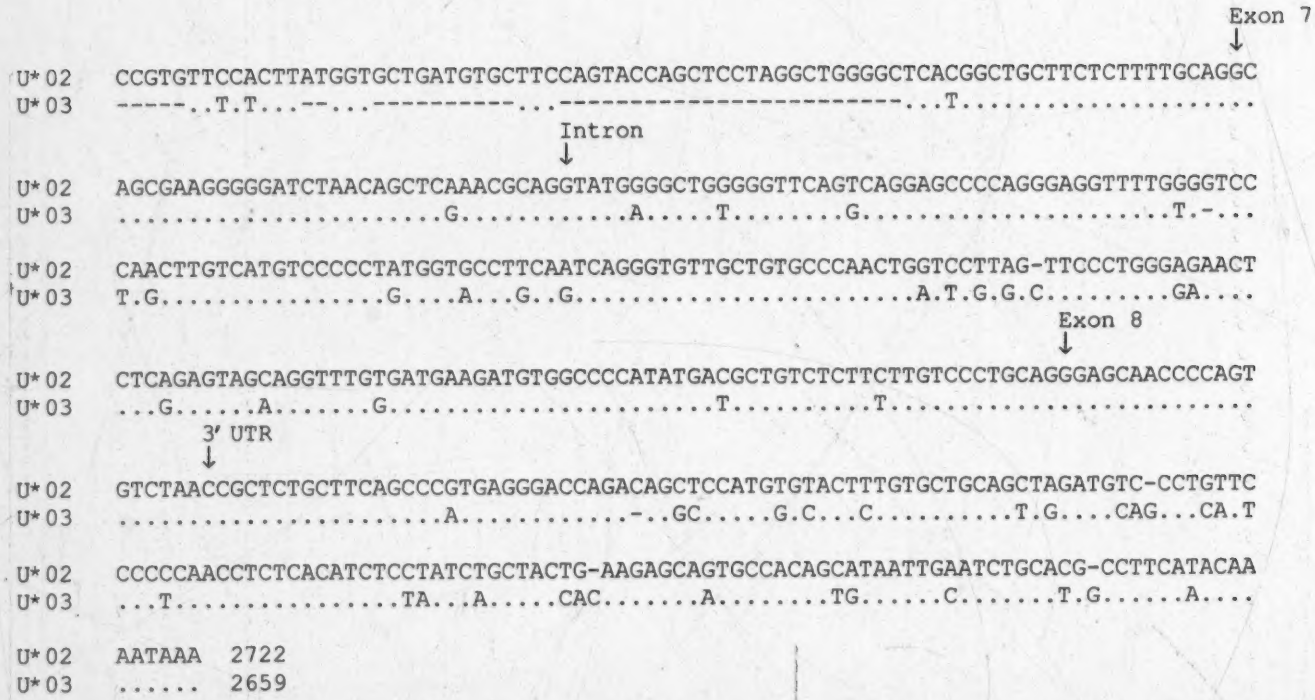


Figure A 7: Genomic nucleotide alignment of the dominantly expressed MHC class I alleles U*02 and U*03, from the positive DNA strand. PCR amplification and sequencing of D26 genomic DNA from exon 1 to exon 6, and from exon 5 to TAP2, allowed a contig of the major locus to be generated. Arrows mark the beginning of introns and exons. Dots indicate identity with the sequence U*02. Dashes are included to maximize identity of the alignment.